Factors influencing Lecturer Assessment Practice in Diverse Southern Contexts

By Danica Anne Sims

(SMTDAN013; 1374964)

SUBMITTED TO THE UNIVERSITY OF CAPE TOWN

In fulfilment of the requirements for the degree

PhD in Health Sciences Education

Department of Health Sciences Education

Faculty of Health Sciences

University of Cape Town

December 2019

Supervisor: Professor Francois Cilliers

Declaration

I, Danica Anne Sims, hereby declare that the work on which this dissertation/thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signed by candidate

Danica Anne Sims

December 2019

Acknowledgements

I have written these acknowledgements countless times in my head, almost on a daily basis, as my PhD slowly began to draw to a close. One skill I have picked up on my journey into the world of education is critical reflection. Often, I would take a moment to press pause and reflect on the many people who have brought me to this point. As they say, I stand on the shoulders of giants. My name may appear on the title page of this thesis, yet a multitude more belong up there.

To my beloved **husband** and best friend, **Raynard Sims**, *we did it*! How can I say thank-you in a single paragraph, when you have been my daily strength and support for the past decade – all of which has been spent in Higher Education! Your unwavering belief in me, especially in my moments of doubt and despair, has carried me through – I could never have come this far without you. Thank-you for your sacrifice and selflessness, working hard to provide for both of us, allowing me the opportunity to study full-time. It has been a long road, with ups and downs, twists and turns, lots of uncertainty and fear, but also so many rich rewards, amazing adventures, learning and growth, abundant joy, laughter and utter fun.

To my **family**, for your love, encouragement and prayers. My parents, **Peter** and **Dagmar Smuts**, tell the story of when my dad was graduating with his PhD (in Theology):

> I was four-years old and my mother told me that my dad was going to become a "doctor". My response was, "So, he is going to make sick people better?" My mom said, "No, he's going to be a doctor of the Bible", to which I replied, "So, he is going to raise people from the dead?"

My parents were my first teachers. They taught me, through their lived examples, to love reading, to work diligently and to the very best of my abilities (always), to be patient, humble, gracious, generous, self-sacrificing and to serve others, equipping me with skills I would need, not only for my studies, but for life. What a blessing and privilege – I will be forever grateful.

My **supervisor**, **Francois Cilliers**. Thank-you for taking a chance and opening the door for me, allowing me the opportunity to step into a new world and pursue my passion. I was enthusiastic but unqualified (and ignorant!), yet, over the last three years, you have helped me build a solid foundation, gently guiding me, but providing me with the space needed to learn, grow and develop into an independent researcher. I will fondly remember our long meetings and interesting discussions, your whiteboard and colourful markers, heavy-hitting and much appreciated feedback and critique, instilling the highest standards of rigour and the need for validated, evidence-based, theory-informed practice. I have thoroughly enjoyed myself.

To the wider **Department of Health Sciences Education** community. I have been challenged, encouraged and inspired by the ongoing work of others in the department. I enjoyed proposal presentations and training seminars, and resulting rich discussions, building friendships and developing comradery with fellow postgraduate students. Thank-you for welcoming me into your tribe.

To my **previous supervisors**, Professor **Sharon Prince** and Dr **Jade Peres**, and the **Prince laboratory**. It was there that I was first exposed to "real" research and "became" (I am still becoming) a researcher. I loved my time in the (largely female) lab and the example of excellence, role-modeling strong women in science. I call myself a "scientist" because of my years under your supervisor and through my interactions with my lab mates and dear friends. I have only been able to build on in my PhD because of what was started by you – thank-you.

To my "PhD friends", who have studied alongside me since our days as first-year science students almost nine years ago! Dr **Jenna Bleloch**, your immense hard work, incredible success and doctoral graduation last year has so encouraged me to keep on going. **Melissa Blumenthal**, you are about to submit your own PhD thesis – the end is in sight! I have greatly admired your quiet grace and resilience. I cannot wait until our "doctoral dinner club" can have a celebratory supper.

To our close family friends, **Peter** and **Tracey Willig**. Your generosity, in providing Ray with a job, has meant that I have been able to study. Without you, we honestly do not how we would have (and continue to) survive. You have given us our livelihood and we do not take it for granted. You have also made us feel like loved family, and that is how we consider you – thank-you from the bottom of our hearts.

To our landlords, **Peter** and **Helen Temple**. Thank-you for your gracious and generous gift in allowing us to stay on your property these last six years. It has been an absolute privilege and pleasure to live in such a beautiful neighborhood and cozy cottage that has become our happy home.

Most importantly, to my **Heavenly Father**. No matter how uncertain, overwhelmed, hopeless, worried, anxious, afraid, doubtful and weak I feel, You are faithful. You are loving. You are generous. You are trustworthy. You provide – you always have and will continue to do so. I cannot take any credit for the abilities you have given me, the parents I have or the home I was raised in, the man I married, the supervisors who led me, or all the wonderful people who have come into my life – it was all according to your plan and goodness. I pray that you will continue to use me as your servant, to give you the honour and praise always. ***Soli deo Gloria***!

Table of Contents

List of Tables and Figures

Glossary

Assessment: situated in a more constructivist paradigm, focuses on the processes of teaching and learning too. Assessment is a broader term, serving many different functions or purposes, to describe both the process of learning (formative assessment, which is ungraded and may be diagnostic) as well as final judgements (summative assessment, which is graded, such as evaluation).

Assessment for learning (AfL): is a positioning of assessment that specifically emphasises a purpose of learning, where assessment is not used to test or measure learning but rather further or increase learning.

Assessment literacy: assessment literacy refers to the ability or capacity of educators to negotiate assessment knowledge and skills in the context of their classroom teaching. It is moderated by context, opportunity to learn, personal preferences and educational culture.

Assessment instrument (or tool or method or format): refers to the type or format of an assessment, such as a written or oral examination, clinical performance or ward-based observation, or portfolio, to name a few (see **Appendix: Table 3.1.** for the major assessment instruments used in HPE).

Assessment of learning (AoL): in contrast to AfL, AoL is a positioning of assessment that emphasises the purpose of assessment to test or measure learning.

Backwash or washback effect: are terms used to describe the effect testing has on teaching and learning; for example, the type of assessment instrument used can influence how students learn for that assessment.

Conceptions: the qualitatively different ways of experiencing, understanding and making meaning of an aspect of reality or the varied interactions and interpretations between a subject (person) and an object (phenomenon) in the real world; depicting a second order perspective of the world.

Consequential assessment: refers to a nuanced understanding of assessment whereby, regardless of whether an assessment is summative or formative, high-stakes or low-stakes, it is perceived as being consequential or having an impact, and influences a student's learning approach or study behaviour.

Critical research paradigm: aim is to emancipate (concerned with social justice), reality is subjective and constructed along power/hierarchies (ontology), truth is many and powerful (epistemology).

Culture: includes values, rituals, heroes, symbols and practice (referring to conventions, habits, traditions, customs).

Epistemology: nature of truth or theory of knowledge.

Evaluation: situated in a more positivist paradigm, focuses on the end product or outcomes to pass a judgement. Evaluation is a narrower term usually used to describe the final marks, grades or judgements post-observation, measuring or testing, and may be a part of an administrative process.

First order perspective: concerned with the object or phenomenon (for example: Phenomenology investigates these perspectives).

Formative assessment: an assessment that is usually unmarked and ungraded, and is a "low-stakes assessment" (see below) as learning is the focus and not judgement.

GLOBE: A large-scale, multi-phase, multi-method research project (950 organisations in 60 countries) focusing on national and organisational culture, with nine said dimensions; performance orientation, assertiveness, future-orientation, humane-orientation, institutional collectivism, in-group collectivism, gender egalitarianism, power distance, uncertainty avoidance. See **Appendix 4.1.** for more details.

Grade: a mark that has been assigned a value or judgement (what that percentage means; for example, a pass, a fail, a distinction, an "A").

Health Behaviour Theory: a theory from the field of (cognitive) psychology (and later sociology) that describes human behaviour as an outworking of various interacting personal and contextual factors.

Health Professionals Education: Health Professionals Education, as opposed to Medical Education, refers to the broader Higher Education of Health Sciences or Health Professionals students; for example, students studying medicine, medical science (basic scientists or laboratory-based/research scientists), health and rehabilitation science (physiotherapy, occupational therapy, speech and language pathology, audiology, optometry), nursing, pharmacy and so on.

High-stakes assessment: an assessment with serious consequences attached to its outcome or result, such as summative assessments where important decisions and judgements are made; for example, passing or graduating from a programme or becoming certified for a particular profession.

Hofstede: A large-scale research project, led by Geert Hofstede, that investigated organisational and national culture; as a result (static) models of culture were developed called Hofstede's Six Dimensions of Model of National Culture (power distance, individualism, masculinity, uncertainty

avoidance, long-term orientation, indulgence) and Organisational Culture (means- versus goal-orientated, internally- versus externally-driven, easy- versus strict-work discipline, local versus professional, open- versus close-system, employee- versus work-orientated). See **Appendix 4.1.** for more details.

Interpretivist research paradigm: aim is to understand and meaningfully describe, reality is subjective and constructed (ontology), truth is many (epistemology).

Learning effects of assessment: are related to the pre-, pure- and post-assessment effects of learning; assessment (before, during and after) is able to influence the quantity and quality of student learning, distribution and level of student effort.

Low-stakes assessment: an assessment with little pressure or consequences attached to its outcomes, such as a routine, formative or ungraded assessment.

Mark: a number-based measurement (such as a percentage).

Maximum variant sampling: a type of purposive sampling, is the selection of respondents from different contexts to increase the likelihood that the findings will reflect differences or varied perspectives.

Measurement: situated in a more positivist paradigm, focuses on providing a judgement based on a measured outcome (usually a number such as a mark, score or grade).

National culture: refers to how groups of nations or regions compare within and without (for example, differences in values).

Ontology: nature of reality or being.

Organisational culture: refers to how members of an organisation relate to one another, their work and the external world. Comparisons are made within and without; inside the organisation versus other organisations).

Paradigm: an interpretive framework, "grand theory" or sets of beliefs and practices that inform how individuals "see" the themselves, the world, knowledge and reality.

Phenomenography: a research methodology that investigates how people experience, interpret and understand what a phenomenon is; the various, diverse, spectrum or range of understandings (a second order perspective).

Phenomenology: a research methodology that investigates what a phenomenon is at its core or consensus or singular essence; the common understanding (a first order perspective).

Positivist research paradigm: aim is to discover natural laws to predict future events, reality is objective (ontology), truth is singular (epistemology).

Programmatic assessment: is a relatively new hybrid design of assessment that attempts to maximise the benefits of formative and summative assessment theory and practice that consists of several (longitudinal or continuous) low-stakes assessments that are used formatively, but also as individual data points that may be combined for a final, summative high-stakes decision-making or judgement. Programmatic assessment may consist of various components including: training or learning tasks, supportive tasks, assessment tasks, learner reflection and planning, social interaction and expert judgement.

Purposive sampling: the selection of participants that have experienced the event of interest or are most qualified to provide rich perspectives, in other words, these individuals meet the predetermined, relevant criteria.

Qualitative research: a scientific method, usually of non-positivist research paradigms and the field of social sciences, to gather non-numerical data, for example, interpreting meaning or developing understanding of a phenomenon.

Quantitative research: a scientific method, traditionally of the positivist research paradigm and field of natural sciences, to gather numerical data and usually consists of statistical and mathematical analysis of numbers and text.

Reflexivity: an ongoing process of critical self-reflection, disclosure and declaration of self, to be aware of and minimise any potential (negative) impact of the researcher on the research process and outputs.

Reliability: refers to the reproducibility of data collected or information gathered, or the consistency of assessment outcomes over time and space; in other words, reliability estimates the amount of random measurement error in assessments (see **Appendix 3.1.** for quality control criteria for designing a technically sound assessment, which includes reliability).

Rigour: criteria for achieving high quality, robust and credible research.

Score: similar to a mark; a number-based measurement.

Second order perspective: concerned with how the subject interprets an object or phenomenon (for example: Phenomenography investigates these perspectives).

Self-regulated or self-directed learning: both self-regulated and self-directed learning speak to the active role and responsibility students themselves play in their own learning processes; students autonomously monitor, guide, update and control their own learning whenever necessary.

Southern Theory: is a (standpoint, post-colonial) theory that conceptually positions knowledge production in the global South in contrast or challenge to the dominant discourses of the global North (or West). It is concerned with critical, emancipatory, socially just research that amplifies marginalised or peripheral voices.

Standpoint Epistemology or Standpoint Theory: is a (critical, empowering and emancipatory) theory that argues that knowledge is situated or positioned (non-neutral) and thus knowledge and power hierarchies exist; various standpoints or positions may be taken in producing knowledge (for example: a feminist standpoint or perspective).

Summative assessment: an assessment with a mark or grade attached, and so usually a "high-stakes assessment" (see below) as decisions, judgements and consequences are attached to theses outcomes or results.

Sustainable assessment: a complement to summative and formative assessment practices with students as active participants being prepared for both current or immediate goals, as well as future assessment activities outside of HEIs (such as post-graduation and in their future careers and professions). Sustainable assessment focuses on testing cognitive goals or cognitive capacities, to ensure the outcome of longer-term learning by students; the goal is a gradual shift from teacher-initiated learning to student-driven learning and informed judgement, where the student takes on a more active and eventually independent role in their learning.

Testing effect or test-enhanced learning: evidence shows that taking tests results in improved learning and performance of students, specifically enhancing long-term retention of knowledge.

Validity: refers to the degree of meaningfulness in the interpretation of assessment outcomes (for example, test scores, not the assessment instrument itself) or the degree to which meaningful conclusions may be drawn because the assessment did indeed assess and deliver according to its intended purpose. Validity is an indication of the level of confidence that the

desired construct was actually measured, supported by several sources of evidence (such as, the content of an assessment, response processes, internal structure of an assessment, relations with other variables measuring the same subject and consequences or impact on learners, teachers and the curriculum) (see **Appendix 3.1.** and **Table 3.1.** for quality control criteria for designing a technically sound assessment, which includes validity).

Abstract

Assessment practice in Health Professionals Education (HPE) has serious consequences for the student and public as it impacts on student learning and outcomes, ultimately certifying a graduate as safe for public practice, and thereby affecting patient care. The goal is for assessment to be practiced in such a way as to drive student learning and outcomes in a desirable manner using assessment to help contribute to the creation of powerful learning environments. Critically, this may not take place without the assessor.

In resource-constrained, Southern contexts, the individual lecturer is usually responsible for practicing assessment, as opposed to a collective assessment committee. It is crucially important to explore how lecturers practice assessment and if their practice positively drives learning. Although lecturers are the key role players in assessment practice in the South, little is known of lecturer HPE assessment practice in the global South. Additionally, these lecturers in HPE generally have no or little formal training in assessment. There is a need for evidence-based, theory-informed, valid and appropriate interventions for faculty training and continued professional development that target lecturer assessment practice.

I propose that lecturer assessment practice is a behaviour, and that how lecturers' think of assessment (their underlying understanding or conceptions, including assessment literacy) and interacting factors (personal and contextual influences), shape their assessment behaviour. In order to explore this behaviour, the conceptual frameworks of Health Behaviour Theory (HBT) and Southern Theory were employed as theoretical underpinnings guide this research study into lecturer assessment practice in the global South. To this end, using purposive and maximum variant sampling, lecturers in diverse Southern contexts were interviewed (South Africa and Mexico) and lecturers' conceptions of assessment and factors influencing their assessment practice were identified and described in a Phenomenographic outcome space and novel HBT-derived model of lecturer assessment practice respectively.

The findings from this study, while needing to undergo further validation in additional Southern contexts, may assist in guiding the design and implementation of strategic and targeted faculty assessment training interventions to enhance lecturer assessment practice leading to the creation of powerful learning environments, thereby improving student outcomes and ultimately improving patient care.

Chapter 1: Orientation

"Assessment drives learning" is a well-accepted phrase in educational literature; yet evidence showing that assessment is able to drive learning in desirable ways remains scarce. It is the goal of educators to use assessment to contribute to positive student learning behaviours and outcomes. This is especially so in the context of Health Professionals Education (HPE), as the consequences of poor assessment practice can have potentially fatal effects on the public. Importantly, any attempts to use assessment to create powerful learning environments may not take place without the assessor – the lecturer who practices assessment.

Currently there is a paucity of information regarding lecturer assessment practice in the context of Health Professionals Education (HPE), especially from the global South. Literature on lecturer assessment practice may focus on a number of areas, including the various roles lecturers play in assessment processes (programme design, item design, appraisal or evaluation) or their assessor behaviour (rater behaviour such as decision-making or feedback practice). In this study, the focus will be on how lecturers understand and practice assessment, that is, why they do what they do.

Based on the context in which lecturers practice assessment (ie: resource-rich or resource-constrained, developed or developing settings), assessment practice may be practiced by collective assessment committees or teams, or the major responsibility may lie with individual lecturers. In the global South, a generally more resource-constrained setting, the assessment role of the individual lecturer becomes more prominent and significant. What these individual lecturers think and understand about assessment, how and why they practice assessment as they do, is important to know, given their role in assessment practice and as an important unit of intervention in any attempts to change assessment practice.

Additionally, in the context of HPE, lecturers who practice assessment generally have no or little training in assessment (absent or poor assessment literacy), declaring the need for assessment

training and faculty development. However, before lecturer assessment behaviour may be changed, it first needs to be understood.

In this thesis I argue that lecturer assessment practice is a behaviour. I employ Health Behaviour Theory (HBT) as a conceptual framework upon which to explore, understand and conceptualise lecturer assessment practice because it provides a rich, theoretical and operationalised approach to explaining and changing behaviour. As this research is limited to resource-constrained contexts in the South, Southern Theory is also used to frame these investigations as it is specifically able to illuminate behaviour in the global South and to position the findings on this study in a global (knowledge) context.

To this end, the research problem centres around lecturers and their assessment practice, and the research questions focus on exploring what factors influence lecturer assessment practice, including a sub-study on lecturers' conceptions of assessment. As stated by HBT, it is important to first understand a behaviour before it may be changed. In this instance, an understanding of current lecturer assessment practice is a prerequisite for identifying and developing theory-based interventions that may then be used for targeted and valid faculty development interventions that potentially lead to assessment driving learning in a positive manner. Southern Theory, stresses the importance of creating contextually-appropriate, relevant and powerful theory in, by and for the global South.

In order to explore these research questions, individual HPE lecturers in diverse Southern contexts (South Africa and Mexico) were interviewed. The end goal of this study is to provide a rigorous and rich description of lecturer assessment practice in diverse Southern contexts, which may then be used as evidence-based platform to both inform the design and development of a sound faculty development intervention to create a powerful learning environments through enhanced lecturer assessment behaviour, as well as contribute to Southern Theory.

To briefly outline the structure of this research project and thesis, **Chapter 1** contains an overview of this study, providing a brief introduction and argument for this study, including the research problem and questions, followed by a layout of the thesis.

Chapter 2 provides the theoretical underpinnings on this research project, describing the conceptual frameworks employed, HBT and Southern Theory. HBT is well suited to the task as it seeks to not only describe but understand, explain and potentially predict behaviour using a developed model including interacting personal and contextual factors. Southern Theory, as a particular type of Standpoint Theory or Epistemology, conceptually positions this study in a global (knowledge) context and moves it beyond mere description.

Chapter 3 provides a detailed literature review of the current research related to lecturer assessment practice, including looking at what others have said about assessment more generally in Higher Education (HE); the importance and consequences of assessment in HPE, briefly touching on the public and student, in particular the learning effects of assessment, before focusing on the personal and contextual factors influencing lecturer assessment practice. A lack of literature and research conducted in the South highlights the gap, need and rationale for this research and the use of Southern Theory as a conceptual framework. In light of the theoretical perspectives and literature reviewed, the research questions will then be detailed.

Chapter 4 describes the Interpretivist methodology and qualitative research approach selected, as well as the use of Phenomenography to investigate lecturers' conceptions of assessment. The purposive and maximum variant sampling strategy and semi-structured qualitative research interviews as the method used for data collection were detailed. The methods of Phenomenographic and Thematic analyses utilised were described and strategies to ensure rigour throughout were mentioned.

Chapter 5 reports on the results of this study, specifically on the first data set collected in South Africa, referred to as the first phase or "pilot study". Each results chapter briefly touches on the

methodology and analyses used, and then sets out to detail the range of conceptions described by the participants; thereby seeking to answer the first research question looking at lecturers' conceptions of assessment.

Chapter 6 continues to report on the results of this study, moving on to the second data set collected in South Africa, and then the final data set collected in Mexico; referred to as the second and third phases or "confirmatory studies". As in Chapter 5, methodological details are given, followed by a description of conceptions.

Chapter 7 seeks to answer the second research question and describes the additional personal and contextual factors that influence assessment practice in diverse Southern settings.

Chapter 8 summarises the findings of this study, and provides final discussions and conclusions, including the potential utility of the results found, limitations of the study and possible future directions for further research into these matters.

At the end of this thesis **References** and additional resources in the **Appendix** are supplied, which includes related background information, participant consent forms, a research reflexivity declaration, and supplementary evidence (participant quotations).

Chapter 2: Theoretical Perspectives

Assessment drives student learning (Schuwirth & van der Vleuten, 2010), and, in the context of HPE, it also impacts on the public in terms of certifying graduates as safe for practice on patients. The practice of assessment needs to be done in such a way as to drive student learning and outcomes in a positive manner. The practice of assessment, however, may not take place without the assessor. What is currently unknown, especially in the global South, is what impact the lecturer has on the practice of assessment? More specifically, how do they practise assessment; what and why?

In the global South, resource-constrained settings in general, HPE assessment is generally practiced by an individual lecturer. Despite the high stakes nature of assessment in HPE, these lecturers too, most often trained clinicians or basic scientists, possess little or no training in assessment and thereby have a poor assessment literacy (Medland, 2018; Popham, 2009; Schiekirka-Schwake, Anders, von Steinbüchel, Becker, & Raupach, 2017).

In fact, there is some evidence that lecturers are under-prepared for assessment practices in HEI, with some authors even referring to assessors as “assessment illiterate” (Medland, 2018; Popham, 2009). For example, studies have reported suboptimal performance of various assessment methods such as global competence ratings (Daelmans et al., 2005; McGill, van der Vleuten, & Clarke, 2013), observation (Holmboe, 2004; Kogan, Conforti, Yamazaki, Iobst, & Holmboe, 2017) and oral assessments (Burchard, Rowland-Morin, Coe, & Garb, 1995), as well as variation and imprecision in assessment scoring (Alexander, Osman, Walling, & Mitchell, 2012). Indeed, an international survey of medical educators revealed that less than a quarter reported having any HPE qualification, despite the majority of HPE educators having spent an average of more than a decade in the field (Huwendiek et al., 2010). Huwendiek et al. (2010) continued to say that medical educators themselves are aware of their need for training and professional development, including improving assessment skills. Moreover, despite perceived general educational expertise, HPE educators expressed a desire for more time to be spent on HPE issues

and to pursue further HPE qualifications (Huwendiek et al., 2010). Davis (2018) recommends faculty development programmes to determine staff competencies and provide educational training in order for the success of medical schools of the future.

This highlights the need for faculty training and professional development interventions; specifically evidence-based and theory-informed, which are currently lacking (Bahar-Ozvaris, Aslan, Sahin-Hodoglugil, & Sayek, 2004; Grossman & Salas, 2011). It has been proposed that the limited success in changing behaviour, such as using assessment to drive learning, may be due to a lack of using theoretical, validated models to guide such interventions (Albert & Reeves, 2010; Bordage, 2009; Cook, Bordage, & Schmidt, 2008; T. Gibbs, Durning, & van der Vleuten, 2011; Godin, Belanger-Gravel, Eccles, & Grimshaw, 2008; Harrison, Konings, Schuwirth, Wass, & van der Vleuten, 2017; Prideaux & Bligh, 2002; Sheeran, 2002; Sorinola, Thistlethwaite, Davies, & Peile, 2017; van der Vleuten & Driessen, 2014). Only a handful have proposed theoretical frameworks involving assessment and learning such as self-regulated learning, but none of these frameworks have been validated (Broekkamp & van Hout-Wolters, 2007; Ross, Green, Salisbury-Glennon, & Tollefson, 2006; Ross, Salisbury-Glennon, Guarino, Reed, & Marshall, 2003; van Etten, Freebern, & Pressley, 1997). Cilliers and colleagues are one of the few who have developed and are validating a theoretical model, describing both the nature and mechanism of how summative assessment impacts learning of medical students in university classroom and clinical contexts (Cilliers, Schuwirth, Adendorff, Herman, & van der Vleuten, 2010; Cilliers, Schuwirth, Herman, Adendorff, & van der Vleuten, 2012).

Before lecturer assessment practice can be changed, it first needs to be understood and explained. I propose that lecturer assessment practice is a behaviour, and so the goal would be to investigate these behaviours, establishing what these lecturers think (their conceptions, including assessment literacy), what they do (their practice), and why they do so (relating to their conceptions and influencing factors), before conceptualising how their assessment behaviour may then be potentially changed. Thus, the individual lecturer who practices assessment in the global South is the unit of analysis in this study, for they are the responsible and key role-players

in HPE assessment in the South and an important unit of intervention in any attempts to change assessment practice. Once their assessment behaviour is understood and explained, these factors and conceptions may be used to potentially inform and specifically design targeted faculty training and continued professional development interventions to change lecturer assessment practice in order for lecturers to practice assessment that drives student learning and outcomes in powerful ways.

Health Behaviour Theory (HBT) was selected as a conceptual framework in which to investigate and conceptualise lecturer assessment practice for it considers the behaviour (and behavioural change) of the individual in context. Southern Theory, relating to Standpoint Theory or Epistemology, was also employed to further conceptualise lecturer assessment practice beyond pragmatic descriptions (such as "resource-constrained" or "developing") and place this study in a global (knowledge) context and to address the need to build theory for the global South, in and by the global South.

2.1. Health Behaviour Theory as a Conceptual Framework for understanding lecturer assessment practice

In seeking to explore the behaviour (assessment practice) of lecturers with the view to ultimately intervene in ways that change behaviour, for example through providing strategic or targeted interventions or training for lecturers that practice assessment, a conceptual framework that explains and predicts behaviour in context was needed (Glanz & Bishop, 2010; Glanz, Rimer, & Viswanath, 2008; Michie et al., 2005). To this end, HBT was employed as an organising framework to study lecturer assessment behaviour in resource-constrained settings.

HBT draws on social and behavioural science theory and is related to behaviour change science; it was initially developed in the field of (cognitive) psychology, later included insights from sociology and more recently has included elements from the fields of anthropology and economics (Bartholomew, Parcel, Kok, Gottlieb, & Fernandez, 2011; R. Davis, Campbell, Hildon,

Hobbs, & Michie, 2015; Kok, 2014; Kok et al., 2016). It seeks to understand the behaviour of individuals, with the *individual as agent*, and later expanded to include the role of the context in behaviour, thereby positioning the *individual in context*. HBT represents a family of theories that seek to describe, explain, predict and change (health) behaviour in, for instance, public health promotion strategies (Bartholomew et al., 2011; R. Davis et al., 2015; Eaton, Flisher, & Aarø, 2003; Glanz & Bishop, 2010; Glanz et al., 2008; Harrison et al., 2017; Kok et al., 2016; Lippke & Ziegelmann, 2008; Michie et al., 2005; Munro, Lewin, Swart, & Volmink, 2007). Several examples of prominent health promotion HBTs that have been adapted and used in HPE include the Theory of Reasoned Action; Theory of Planned Behaviour; Integrated Behavioural Model; Health Belief Model; Trans-theoretical Model and PRECEDE-PROCEDE model; yet, as this study is not interested in a specific HBT, for no single HBT provides a complete explanation of a behaviour and causal elements and constructs are shared within the range of HBTs (with unique or nuanced factors usually relating to the specific health behaviour the model is seeking to address), however, various HBTs within the family of theories were consulted and a more general and eclectic HBT model was employed in this study that can be adapted to numerous behaviours and contexts (Bartholomew et al., 2011; Cilliers, Schuwirth, & van der Vleuten, 2015; Eaton et al., 2003; Lippke & Ziegelmann, 2008; Munro et al., 2007).

HBT possesses a rich tradition of theory building and operationalised (applied) research, which has more recently been employed in Higher Education (HE) and HPE research. By the way of an example, while the Theory of Planned Behaviour may have originally been intended for health promotion, it has been used in HPE to teach professionalism, specifically addressing student attitudes, social norms and behavioural intention (Archer, Elder, Hustedde, Milam, & Joyce, 2008; Rees & Knight, 2007). There is a precedent for using HBT to describe and explain (and potentially ultimately change) lecturer assessment practice for it has already been used to describe and explain HPE student learning behaviour (Cilliers et al., 2010; Cilliers, Schuwirth, Herman, et al., 2012; Cilliers, Schuwirth, & van der Vleuten, 2012a, 2012b) and proposed as a theoretical position for exploring and designing interventions for faculty development and training in HPE (Cilliers et al., 2015). Overall, HBT provides a theoretical framework within which to approach the complex

behaviour of lecturer assessment practice in an ordered manner and allows for the design of conceptually strong research.

In describing the commonalities across the family of HBTs, HBT proposes that behaviour is the complex interaction between internal or personal and external or contextual factors (Cilliers et al., 2015; Glanz et al., 2008). Contextual factors may be divided into proximal and distal contextual factors. Proximal factors may be subdivided into interpersonal (interaction with others), physical (environment, availability and access to resources and infrastructure) and organisational (the company or institution within which an individual may be). Distal factors include structural (legal, political, economic, social factors) and cultural (shared values, beliefs, norms, traditions, discourse and variations) factors. Personal factors may be divided into intrapersonal (knowledge, conceptions, self-efficacy, self-esteem) and behavioural (actions) (see **Figure 2.1.**) (Cilliers et al., 2015; Eaton et al., 2003; Maxwell, 2004a; Michie et al., 2005).

Critically, these different factors interact with one another to influence the final action or behaviour and thus must be considered together. This is summarized well in the following quote, "Individuals exist within groups, which, in turn, are embedded within organisations and higher-order systems. The individual is influenced by these systems and can, in turn, influence them directly or through groups and organisations" (Bartholomew et al., 2011). This is important because, if an intervention to change a behaviour is directed at a particular level, the multiple other levels must also be considered, for they too may influence the targeted behaviour or the intervention may impact on them (Bartholomew et al., 2011). A strength of HBT is that is considers the individual in context.

While HBT describes lecturer assessment practice in context (literature describing personal and contextual factors that impact on lecturer assessment practice will be detailed in **Chapter 3**), this research was positioned within the global (knowledge) context too, through the use of Southern Theory as an additional conceptual framework to illuminate lecturer assessment practice in the South.

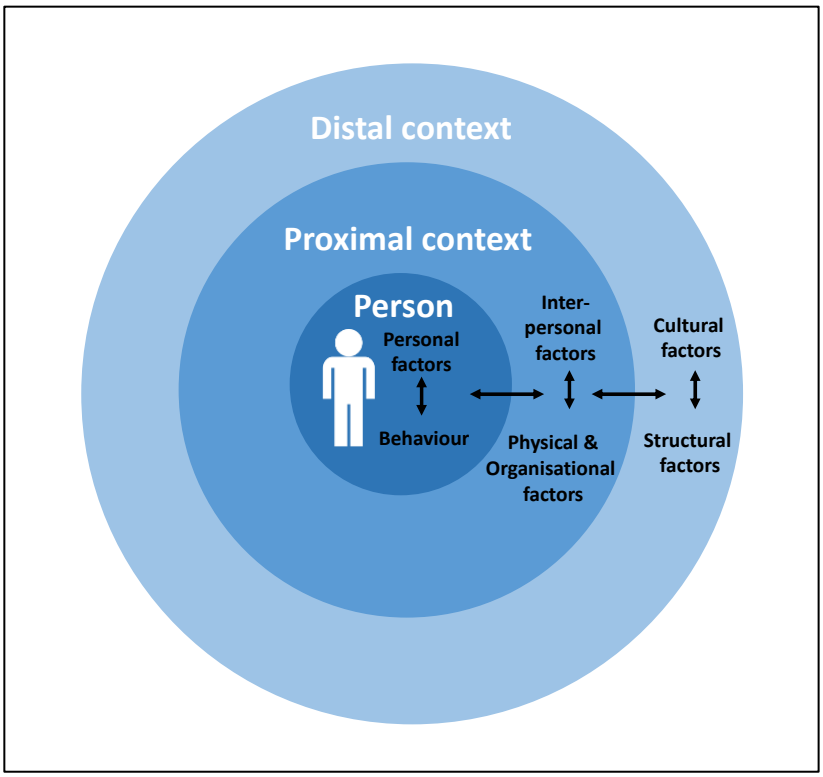


Figure 2.1.: Health Behaviour Theory: an overview. A general example of a Health Behaviour Theory (HBT) framework of an "embedded reciprocal system" for organising the relationship between factors at various levels that impact on behaviour (Bartholomew et al., 2011). Adapted from Eaton et al. (2003).

HBT was chosen as a conceptual framework in which to explore lecturer assessment practice. A generic, eclectic and adaptable HBT model, broadly consisting of personal and contextual factors, organised in an "embedded reciprocal system", was selected for this study (Bartholomew et al., 2011). HBT provides an easily accessible theoretical perspective to study behaviour and behaviour change. A generic model allows for both deductive and inductive analysis, and, powerfully, considers the individual within their context. Furthermore, as HBT has a rich history of behaviour change, this conceptual framework lays the foundation for future professional development of lecturers who practice assessment (Cilliers et al., 2015).

2.2. Southern Theory as a theoretical perspective to illuminate lecturer assessment practice in the global South

In considering the development of theory and practice, such as interventions, training or professional development programmes, these need to be contextually relevant and appropriate. To help ensure applicability to any given context, one must ask the following questions of any account from another context: where was this theory or intervention produced, by whom and for whom?

Generally, there is a mismatch between where knowledge is produced and where it is consumed (Connell, Pearse, Collyer, Maia, & Morrell, 2018a, 2018b). Indeed, it has been argued that most behavioural research is done in Western educated, industrialized, rich and democratic societies which are poor representations or generalisations to the majority of the world's population (Henrich, Heine, & Norenzayan, 2010).

This mismatch has been documented in HPE too; where "global" HPE standards (graduate competencies, definitions of professionalism, teaching, learning and assessment methods) are assumed to be applicable and are adopted in non-Western settings (Gosselin, Norris, & Ho, 2016). This is problematic because no consideration of the difference in cultures are taken into account. Gosselin et al. (2016) describe this phenomenon as "Western cultural imperialism." HPE research needs to be contextually and culturally appropriate to increase application and sustainability, and to limit resource wastage and unnecessary, costly failure (Gandomkar & Sandars, 2017; Walubo et al., 2003).

Furthermore, a study reported the gross under-representation of medical education publications from researchers in Sub-Saharan Africa: a total of 3749 articles during the 45-year time period of 1965-2010 (Greysen et al., 2011), compared to the 2953 articles that were published, from just two medical education journals (Medical Teacher and Academic Medicine) during a 5-year time period (1995-2000), with 74-95% of authorship coming from only five countries: USA, Canada,

UK, Netherlands and Australia (Tutarel, 2002). This finding is supported by more recent literature; a handful of developed countries dominate medical education research and publishing (Doja, Horsley, & Sampson, 2014). This too was observed in unequal funding and the quality of research outputs (Reed et al., 2007).

This skewed geographic distribution is sometimes referred to as the North-South gap, with developed and well-resourced countries residing in the global North compared to the developing and resource-constrained countries in the global South (Greysen et al., 2011). This North-South gap[1] matters for both practical and theoretical reasons, practically because the context in which HPE is practiced, specifically in resource-constrained settings, differs significantly from where the bulk of the HPE literature emanates (Gosselin et al., 2016; Henrich et al., 2010; Rotgans, 2011; Tutarel, 2002; Walubo et al., 2003), and, theoretically, because, as argued in Standpoint Theory or Standpoint Epistemology, knowledge production is not neutral but socially positioned or situated (Fawcett & Hearn, 2004).

Standpoint Theorists argue that, because knowledge is socially positioned and powerful, alternative perspectives, such as those from the margins of "otherness" need to be considered and researched in order to obtain an un- or less-biased version of knowledge or the truth; an understanding that is not influenced or informed by a social (hierarchical) model (Fawcett & Hearn, 2004). Harding (2009, p. 195) describes this knowledge as being less partial and less distorted by social position as it is a more, "Accurate, comprehensive, objective, and rational production of knowledge." Standpoint Theory acts in contrast to the hegemonic masculinist positioning of knowledge (Stoetzler & Yuval-Davis, 2002). One study argued that a "collective hermeneutic gap" could typify the North-South division, whereby the North, with its dominant collective understanding, fails to understand the "social experience of the powerless (global

[1] Importantly, when referring to North and South, it speaks to both the context *and* the content of knowledge production. Importantly, the North-South divide is porous and constantly shifting, but generally the North refers largely to Europe and North America and the South to Africa, Central and South America, Asia and Australia (to some extent) (Comaroff & Comaroff, 2012).

South)" and it would be inappropriate to use theory and practice emanating from North in the South (Luckett, 2016). This hermeneutical gap has been described as an unavailable or inaccessible viewpoint: "In hierarchically organized societies, the daily activities and experiences of oppressed groups enable insights about how the society functions that are not available – or at least not easily available – from the perspective of dominant group activity" (Harding, 2009, p. 194). Southern Theorists Connell, Collyer, Maia, and Morrell (2017, p. 26) demonstrate this through using citation patterns: "We find, empirically, (knowledge) workers in peripheral countries primarily citing the texts of authors from the global North, while workers in the North mostly cite each other, and mostly ignore the ideas and studies produced by workers in the global South. The overall effect has been a structuring of knowledge production where Northern-produced knowledge is treated as the "gold standard", while the possibilities for disciplinary diversity and innovation are constrained."

This is demonstrated in other ways, doctoral students from the global South, studying in the United Kingdom, researching internationalisation, produced some new and innovative perspectives, potentially contributing to Southern Theory, yet also adhered to and reproduced the dominant discourse of existing and well-established Western perspectives: "The theories and methodologies adopted by the students indicate that PhD students, particularly, those from the global South, are strongly influenced into developing Northern theory and using Western lenses" which somewhat reinforces long standing global divisions and patterns of power inequalities (Montgomery, 2019, p. 134). This displays the social power and positioning of knowledge, the valuing of the dominant, Western or global North theories and methodologies and the difficulty in orientating away from the centre to the periphery or margins to produce alternative theories for the South (Montgomery, 2019).

Similarly, a series of studies conducted by Connell et al. (2016; 2017; 2018) revealed Southern-tier knowledge workers (professionals, researchers, academics) adopt "extraversion" Northern research paradigms (theories and methodologies) for many complex reasons. These include resource constraints, difficulty in setting up a workforce, career sustainability, pressure to publish

in international journals, language, asymmetrical North-South research partnerships, recognition and prestige, and beliefs that Europe and America are the sites of the most accepted and advanced research methods and knowledges (Connell et al., 2018a, 2018b). Furthermore, organisational or institutional pressures for knowledge workers can result in, "(A) conservative repetition of disciplinary norms… developed in the North which appear simply as definitions of best practice, scientificity, or modernity" which reinstates the centrality and dominance of the global North, underlining the need for alternative perspectives or standpoints to destabilize these global hierarchies of knowledge (Connell et al., 2017, p. 26).

Moreover, in a literature review of three influential HPE journals (Academic Teacher, Medical Teacher and Academic Medicine) on non-Western researchers' articles on globalisation in Medical Education, it was found that Western voices and paradigms were dominant and non-Western perspectives marginalised (Gosselin et al., 2016). In terms of globalization and culture the authors found themes of cultural homogenisation (striving towards Western culture as the norm and ideal in terms of training, tools, measurements and language), polarisation (focus on cultural distinctions and the local context) and hybridization (adaption of Western paradigms, belief that the East has paradigms of value to teach the West, and building international, cross-cultural, mutually-beneficial partnerships) (Gosselin et al., 2016). The authors concluded that the goal of global HPE should be to work towards more, "Equitable, context-sensitive and locally-driven approaches to medical education" (Gosselin et al., 2016, p. 691).

This phenomenon has also been exhibited in sub-Saharan Africa. In an analysis of all HPE publications from sub-Saharan Africa, it was found that there has been an expansion of HPE research in both traditional and non-traditional databases (referred to as "grey literature" over time (Greysen et al., 2011). This was accompanied by a skewing of research topics, representation (a disproportional representation of South Africa, Nigeria and Uganda in terms of published authors compared to the rest of sub-Saharan Africa) and a propensity to publish in English and in journals from South Africa, the United Kingdom and United States of America, suggesting persistent coloniality as evidenced by the elevation of the value or importance of the West or

global North (Greysen et al., 2011). In fact, recent debate in South Africa over "decolonising" HE[2] underscores this argument (Ndlovu-Gatsheni, 2013). In line with the decoloniality movement, these questions must be asked: who is producing knowledge, for whom and where? (Ndlovu-Gatsheni, 2013). There is a need not only for the redistribution and transformation of knowledge but also the dismantling, reframing and repositioning of it; a need for cognitive justice, so that Southern academics and their research may be recognised and affirmed (le Roux, 2018; Leibowitz, 2017; Luckett, 2016; Shay & Peseta, 2016).

Assessment, in particular, is argued to be a social practice, for it is dependent on who is designing and executing it (Shay, 2004). Assessment practice is neither value-neutral nor culture-free and can be deeply problematic. "(An) assessment system in which the predominantly white, male, middle-class, Western values of objectivity and individualism are lauded as markers of good work" (Hanesworth, Bracken, & Elkington, 2018, p. 98), following or implementing hegemonic theory and practices originating from the global North, may well be inappropriate in its application and unsuitable for the differing context (Tsai, Ho, Hirsh, & Kern, 2012; Walubo et al., 2003). It may also be a form of oppression in re/producing inequalities and injustices, (Kester, 2018; Luckett, 2016; Shay & Peseta, 2016; Varpio, Ajjawi, Monrouxe, O'Brien, & Rees, 2017).

As knowledge is socially sourced, and society exists within the context of social power, Standpoint epistemology is concerned with the emancipation of knowledge and power (Fawcett & Hearn, 2004). The idea of emancipatory research is an important one, for if all knowledge is situated or positioned, then whether it is generated from the dominant "colonized" centre or oppressed margins, it can never be "innocent." However, if the concern is emancipation, or a contrasting

[2] In 2015, 2016 and 2017, students, across South Africa, and further afield, rose up to protest against several issues in Higher Education in South Africa, including the privileging of Western and "white" ways of knowing, doing and being (under the banner of #RhodesMustFall, referring to the colonial figure of Cecil John Rhodes, who represented alienating and oppressive coloniality in Higher Education in South Africa, yet, also referred to an actual statue of Rhodes, seated in a place of prominence and dominance at the University of Cape Town, which was later removed), and the yearly increasing of HEI tuition (under the banner #FeesMustFall), making HE even more inaccessible to the majority of South African, and causing historic debt and academic exclusion, mostly of students of colour.

and countering of viewpoints, then it achieves its purpose of overcoming social constructs of "otherness" and its agenda of empowerment (Fawcett & Hearn, 2004; Harding, 2009). A goal of Standpoint Theory research is to uncover the hidden aspects of social and power relations, as well as the structures that support these relations (Rolin, 2009). It demands involvement, engagement, participation and collaboration, calling social hierarchies, epistemological dogmas, advantages, privileges and deeply held Western or "Northern" assumptions into question, commits to the "other" and strives for freedom, equality, equity and empowerment and the performance of socially just research: "It is one thing to gesture toward including the excluded in our thinking and social projects. It is quite another to engage seriously not only with their ways of understanding themselves and their social relations, but also with their ways of understanding us and our social relations" (Harding, 2009, p. 193). Southern Theory, as a branch or type of Standpoint Theory[3], offers a lens or conceptual framework within which to interrogate this state of affairs (see **Figure 2.2.**).

Southern Theory purports that there is a global sociology of knowledge and, more recently, there have been post-colonial debates and social perspective changes around knowledge production and intellectual practices. In other words, using Standpoint Theory terminology, the "other", in this instance the global South, has alternative, valid and critical standpoints of knowledge to offer the dominant, global North (Connell et al., 2017). Comaroff and Comaroff (2012, p. 113) outlined the knowledge production problem as, "(The) global South has become a shorthand for the world of non-European, postcolonial peoples. Synonymous with uncertain development, unorthodox economies, failed states, and nations fraught with corruption, poverty, and strife, it is that half of the world about which the "global North" spins theories. Rarely is it seen as a source of theory and explanation for world historical events." The authors, and others, continue to say that the

[3] I am arguing that Southern Theory is a branch or type of Standpoint Theory because of similar underlying concepts and principles. "Southern" is a standpoint and both are concerned with critical, emancipatory social research, amplifying alternative and indigenous knowledges. In fact, "Indigenous Epistemology" and "Indigenous Standpoint Theory" have been conducted in the context of Aboriginal research in Australia which has a history of coloniality (Foley, 2003, 2006).

global South is viewed as a reservoir of raw fact that the global North is able to use, shape and refine for their own purposes, and not as credible knowledge producers in and of themselves (Comaroff & Comaroff, 2012; Morrell, 2016).

Post-colonial Southern Theory seeks to challenge, “Western capitalist modernity… To move the project of theory-making to an “ex-centric site”, thus to capture the restless, re-visionary energy that comes from the vast reaches of the planetary population whose genealogies do not reach back directly into the European Enlightenment… (And) has kept the non-European other ‘in the ‘‘shadow’’ of the Western ‘‘Self’’’ – thereby allowing the Universal Subject to remain securely on Euro-American terrain” (Comaroff & Comaroff, 2012, p. 115). The authors continue to speak of an Afro-modernity, not where similarities are described between knowledge production in Africa and Euro-America, but where Africa brings something new, unique and just as relevant to the table of theory-making (Comaroff & Comaroff, 2012). The purpose of Southern Theory, then, is to challenge the dominance and dependency on the West and North, and to be “Emancipatory, contributing to the democratisation of knowledge,” thereby re-centering the South as a knowledge producer and not just an object of study (Connell, 2014; Morrell, 2016).

In this study, Southern Theory provides the conceptual grounding and justification for the inclusion in our sampling strategy in terms of sampling specifically from a diverse range of Southern sites (South Africa and Mexico). It also provides an emancipatory, global South-driven research approach to conduct research in new directions and to produce a collective and collaborative knowledge production on HPE assessment practices from the periphery.

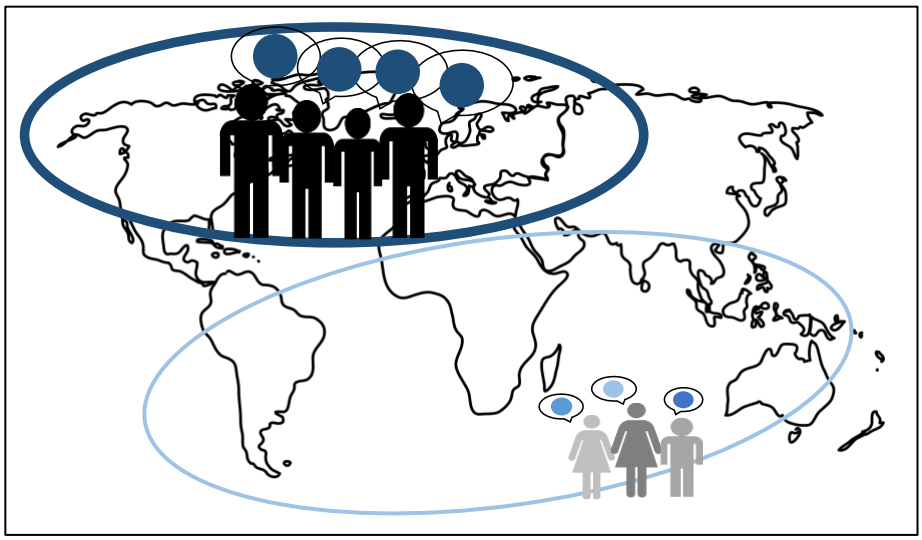


Figure 2.2.: Standpoint Epistemology and Southern Theory: an overview. Standpoint Epistemology purports that knowledge is not neutral but rather is positioned. Positioning implies structures or hierarchies. Standpoint Theory strives to present and empower the marginalised or peripheral voices relative to the dominant positions and discourses. Southern Theory declares that the "global South" is a marginalised knowledge position or perspective in contrast to the "global North" (sometimes referred to as "the West" or "colonial" or "white" or "male" or "hegemonic" positions) and seeks for the Southern voice to be amplified, contributing to knowledge. Importantly, these dominant discourses may not be appropriate or suitable, practically and theoretically, to "othered" contexts and cultures; in fact, they may be problematic and oppressive. Thus, in this research study, Southern Theory provides the theoretical perspective for a sampling strategy centred on Southern, "developing" and "resource-constrained" contexts.

The use of Southern Theory as a theoretical perspective is important for it elevates the research beyond merely providing a pragmatic description of lecturer assessment practice in the South ("resource-constrained" and "developing" contexts). It situates the findings of this study in a global (knowledge) context and thus contributes to Southern Theory through the addition of diversity to the assessment discourse: "A concept of socially just assessment praxis… (and to enhance) educational experiences and learning outcomes for students that have been systematically marginalised by the normative procedural practices that have traditionally informed the nature of supposedly objective assessment" (Hanesworth et al., 2018). This may then lead to the creation of contextually powerful learning environments. Moreover, in performing critical or socially just research, it may be emancipatory and empowering for Southern academics to now have access to powerful and valued othered, alternative and

indigenous knowledges, and further enable their engagement and full participation in society and academia (Shay & Peseta, 2016).

A possible criticism regarding the use of HBT, as a family of theories largely developed in the global North, needs to be addressed in light of Southern Theory. I am aware of the paradox of adopting "Northern theory" as a conceptual framework in the global South. However, since HBT seeks to understand the behaviour of individuals in their context, the Southern context will be considered in exploring lecturer assessment practice. Indeed, I have elected to use a general and eclectic HBT that includes diverse and adaptable personal and contextual elements (beliefs, perceived norms, culture, national politics) and so I do not expect a conflict between the two conceptual frameworks used in this study.

To conclude, this chapter has detailed and provided a rationale for employing HBT and Southern Theory as conceptual frameworks for exploring lecturer assessment practice in the global South. In the next chapter, literature is reviewed to determine what is already known about assessment and how it drives learning, as well as what factors impact on the lecturer assessment practice.

Chapter 3: Literature Review

The focus of this study is lecturers' assessment practice, not simply describing but understanding and explaining their behaviour, by investigating what personal and contextual factors may interact and influence their assessment practice, in line with the chosen conceptual framework of HBT. Southern Theory, as a second conceptual framework employed in this study, is critical as it further illuminates and positions lecturer assessment practice in a global (knowledge) context.

The context in which assessment is practiced needs to be considered, not only because of Southern Theory, but also for pragmatic or logistical reasons, as the role/s the assessor, or lecturer, plays in the assessment process (its design, data collection, appraisal and decision-making) varies from context to context. In general, in the resource-rich global North, assessment in HPE is undertaken by a collective team or committee, whereas in the resource-constrained global South the individual lecturer is often solely responsible for the practice of assessment. As the role of the individual lecturer in assessment in Southern contexts is more prominent and significant, it is essential to know what they think and do, and why, when it comes to assessment.

It is important to understand how and why lecturers practice assessment in the way in which they do because the practice of assessment in HPE has serious consequences for both the student and the public (Boud, 1995), in terms of impacting on student learning and outcomes, and certifying a graduate as safe for clinical practice, ultimately effecting patient care and public safety. Norcini et al. (2011, p. 206) states that, "Assessment in medical education is multifaceted. It drives and stimulates learning, provides information on educational efficacy to institutions and teachers, and protects patients". If assessment is done poorly the consequences can be severe, ranging from patient harm to death (Bordage, Meguerditchian, & Tamblyn, 2013; Makary & Daniel, 2016). There is little literature describing medical error in the Southern contexts but in the USA largely preventable medical error is the third leading cause of death (Makary & Daniel, 2016). This raises the following questions: how may assessment be used to create powerful learning environments, thereby improving student learning and, ultimately, patient care?

I postulate that there are two components to successfully achieving the goal of practicing assessment to drive learning in desirable manners: it first needs to be understood how assessment influences student learning, a study already undertaken by Cilliers and colleagues (Cilliers et al., 2010; Cilliers, Schuwirth, Herman, et al., 2012; Cilliers, Schuwirth, et al., 2012b), and, secondly, we need to understand how assessors practice assessment and what influences their assessment practice. Indeed, any intervention to use assessment to drive learning cannot occur without the faculty involved, the aim of this research project is to explore lecturers' practice of assessment and the factors that influence it.

3.1. Literature Review Approach

The construction of the literature review in this thesis was informed by five research review traditions as described by McGaghie (2015) . These traditions are: narrative, systemic, scoping, critical-realist and open peer-commentary (McGaghie, 2015). A critical-realist research review strategy was adopted, which is a hybrid between the narrative, systematic and scoping traditions (McGaghie, 2015). This approach includes the following steps: (1) Scoping the review, specifically around the research question, review purpose and theories consulted; (2) searching the literature based on relevant evidence for the first step; (3) synthesizing the findings and; (4) developing a narrative (McGaghie, 2015).

The research problem was defined (lecturer assessment practice), the purpose clarified (improving critically-important HPE assessment) and theories to be used identified (HBT and Southern Theory). Search terms and the sampling strategy were guided by the theoretical frameworks employed.

Initial search strategies included searches of "assessment" with variations of "higher education", "medical education", "health sciences education" and "health professions education" on Google Scholar and EBSCO. I would make a note of citation numbers in order to determine the extent of engagement with the paper.

Narrower searches, such as “assessment for learning” or “programmatic assessment” took place after the initial, broad searches, to specifically address points of interest in the arguments being made. This nonlinear and interactive search strategy was in agreement with scoping reviews. This approach was also related to a narrative traditions, for literature was particularly searched for in order to further develop the story being told. The fact that a variety of literature was consulted supports the critical-realist tradition, as it revealed the general landscape of the field, with its various perspectives, highlighted the gaps, and, the appropriate research problems and questions.

As the research project developed, additional concepts were specifically explored, for example, “conceptions”, “conceptions of teaching”, “conceptions of learning”, “conceptions of assessment”, “factors”, “personal factors”, “contextual factors” and “identity”, “professional identity”, “identity formation”, “identity development”, with variations of “higher education”, “medical education”, “health sciences education” and “health professions education”.

These more specific searches speak to the core of the critical-realist approach, searching for literature related to the defined research problem and questions (lecturers’ conceptions of assessment, and personal factors influencing assessment practice). These particular searches were also related to the theoretical frameworks consulted in this study. For instance, as Southern Theory was used to theoretically justify the sampling strategy, so commentary on research generally, and HPE specifically, in the global South was searched for.

In terms of the inclusion criteria, according to a critical-realist review tradition, literature was selected based on its compelling and satisfying content. The literature reviewed was not exhaustive, nor absolutely systematic, but, rather, relevant to the narrative being told and rigorous in its evidence in support of the arguments made. Indeed, findings from different studies were qualitatively compared and contrasted, and, confirmatory and contradictory studies were

searched for. Furthermore, the literature included was coherent with the theoretical frameworks used.

Throughout the iterative literature review search, a record of the papers read and reviewed were noted in my PhD research journal (totally 150 000 words). This journal transparently tracked the development of the narrative and arguments made as the research progressed.

The goal of a critical-realist review tradition, is to write a convincing narrative, synthesizing and integrating complex findings, that also seek to further develop and move from theory to practical implications and interventions; which is in line with the ultimate outworking of this thesis (lecturer training and improved assessment practice).

3.2. Assessment in HPE

In general, the procedures and process of assessment, and the different roles lecturers play, include its design, the collection of data or information, its appraisal and the decision-making tied to the data collected from the assessments. In line with these procedures, assessment may be defined as, "Any *systematic methods* used to *obtain information* about people, objects or programmes" (*emphasis* added) (Downing & Yudkowsky, 2009). More specifically, in HE, Popham (2009, p. 8) states that, "The fundamental function of educational assessment… (is) the *collection of evidence* from which *inferences* can be made about students' *skills, knowledge, and affect* (attitude)" and assessment in HPE is, "(Any) action to obtain information about the *competence* and *performance* of a candidate" (*emphasis* added) (Schuwirth & van der Vleuten, 2011a).

3.2.1. The purposes of assessment

Data is collected using assessment for a variety of purposes, which may also relate to how and why lecturers practice assessment. In education, in general, these purposes have been described by Natriello (1987) as selection, direction, motivation and certification (see **Table 3.1.**). (1) Selection refers to gaining entry into or through a programme; (2) direction provides information as to what is presently occurring, what needs to happen next and how to get there; (3) motivation helps drive learning; and (4) certification provides final judgement and approval for exit from a programme (Natriello, 1987). Other authors may have condensed these purposes of assessment to summative and formative (Carless, 2007; Natriello, 1987; Norton, Norton, & Shannon, 2013; Schuwirth & van der Vleuten, 2010), referring to Natriello's (1987) selection and certification, and direction and motivation, respectively, or the related movements of "Assessment *of* Learning" (AoL) and "Assessment *for* Learning" (AfL) (Bienstock et al., 2007; Black, Harrison, Lee, Marshall, & Wiliam, 2003; Black & Wiliam, 1998; Bottomley & Denny, 2011; Crooks, 2011; Dochy, Segers, Gijbels, & Struyven, 2007; G. Gibbs & Simpson, 2004; Isaacs, Zara, Herbert, Coombs, & Smith, 2013; McKimm, 2009; Rushton, 2005; Schuwirth & van der Vleuten, 2011a), or even used other terms such as conventional, educational or competency and authentic assessment (Boud, 1995). Yet, Natriello's four purposes of assessment in education generally encompass subsequent proposed definitions and purposes of assessment.

Related to the various purposes of assessment is the phrase "assessment drives learning", or the learning effects of assessment; indeed, Boud (1995, p. 2) writes, "Assessment always leads to learning. But the fundamental question is, "What kind of learning?""

Table 3.1.: An overview of the purposes of assessment, and related terms, in HPE.

Selection	Direction	Motivation	Certification	References
Selection (entrance and progression)	Direction (diagnosis and direction)	Motivation (drives learning)	Certification (final judgement and exit)	Natriello, 1987
Summative (decision-making and judgement)			**Summative**	• Carless, 2007 • Natriello, 1987 • Norcini et al., 2011 • Norton et al., 2013 • Schuwirth & van der Vleuten, 2010
Assessment of learning (measurement)			**Assessment of learning**	Black & William, 1998
	Formative (learning or improvement)	**Formative**		• Carless, 2007 • Natriello, 1987 • Norcini et al., 2011 • Norton et al., 2013 • Schuwirth & van der Vleuten, 2010
	Assessment for learning (diagnosis and direction)	**Assessment for learning**		Black & William, 1998
Conventional assessment (traditional quantifying of learning)	**Educational assessment** (psychometric measurements)		**Competency and authentic assessment** (post-graduate quality and validity, related to outcomes)	Boud, 1995

3.2.2. The learning effects of assessment

There is evidence that assessment is able to influence student learning behaviours. A helpful organising framework of viewing the learning effects of assessment is given by Dochy et al. (2007) as the pre-, pure- and post-assessment effects as they encompass the range of effects described below (see **Table 3.2.**). Pre-assessment effects of learning occur before an assessment and include the backwash or washback effects of assessment, whereby students adopt varied study behaviours based on their perceptions of the upcoming assessment (Alderson & Wall, 1992; Cheng, 2000; Newble, 2016; Watkins, Dahlin, & Ekholm, 2005). Pure-assessment effects take place during the assessment experience and include recall, application, integration of knowledge, reflection. Post-assessment effects refer to the after-effects of an assessment, such as the testing effect or test-enhanced learning, in which testing improves long-term storage and recall of information in subsequent tests (Agarwal, Karpicke, Kang, Roediger, & McDermott, 2008; Butler & Roediger, 2007; Karpicke, Butler, & Roediger, 2009; Karpicke & Roediger, 2008; Larsen, Butler, & Roediger, 2008; Roediger & Karpicke, 2006a, 2006b). These studies provide support for the phrase that "assessment drives learning" in terms of the quantity and distribution of student effort, quality and level of effort (Dochy et al., 2007) which has been shown in the context of HPE, in both the classroom and clinical settings (Cilliers et al., 2010; Cilliers, Schuwirth, Herman, et al., 2012; Cilliers, Schuwirth, et al., 2012a, 2012b).

These learning effects of assessment clearly demonstrate that assessment is able to influence student learning; yet the question remains; what is the direction of that influence and are lecturers able to use assessment in different ways to direct different components of student learning?

Table 3.2.: An overview of the learning effects of assessment, and examples, in HPE.

Pre-assessment effect	Pure-assessment effect	Post-assessment effect	References
Changes in learning take place **before** an assessment	Changes in learning take place **during** an assessment	Changes in learning take place **after** an assessment	Dochy et al; 2007
Backwash or washback effect: students adapt their learning strategies strategy based on the format of an assessment and its perceived demands			• Alderson & Wall; 1992 • Cheng, 2000; • Newble, 2016; • Watkins et al., 2005
	Testing effect or test-enhanced learning: the practice of retrieval of stored information		• Agarwal et al., 2008; • Butler & Roediger, 2008; • Roediger & Karpicke, 2006a; 2006b
		Testing effect or test-enhanced learning: testing improves long-term retention and recall	• Agarwal et al., 2008; • Butler & Roediger, 2008; • Roediger & Karpicke, 2006a; 2006b

As it is the desire of lecturers to use assessment to drive student learning in a positive direction, contributing to the development of lifelong learners, numerous theoretical and practical approaches have been suggested, including the inclusion of formative assessment ideas (Barrett et al., 2016; Massie & Ali, 2016; Pugh, Desjardins, & Eva, 2017); the specific focus on the use of feedback (Archer, 2010; Boud & Soler, 2016; Carless, 2007; Dochy et al., 2007; Nicol & Macfarlane-Dick, 2006; Rushton, 2005; Yang & Carless, 2013); sustainable assessment (Beck, Skinner, & Schwabrow, 2013; Boud, 2000) and programmatic assessment (Bok et al., 2013; Heeneman, Oudkerk Pool, Schuwirth, van der Vleuten, & Driessen, 2015; Homer, Fuller, & Pell, 2017; Jessop, El Hakim, & Gibbs, 2013; Schuwirth & van der Vleuten, 2011b; Torre, Schuwirth, &

van der Vleuten, 2019; van der Vleuten & Dannefer, 2012; van der Vleuten et al., 2012; van der Vleuten, Schuwirth, Driessen, Govaerts, & Heeneman, 2014; Watling & Ginsburg, 2018).

While there have been a handful of reports suggesting beneficial gains in student learning or performance through sustainable and programmatic assessment approaches (Beck et al., 2013; Driessen, van Tartwijk, Govaerts, Teunissen, & van der Vleuten, 2012), unfortunately, the majority of studies have not delivered as much as was hoped, reporting mixed results, both positive and negative, on the effects of feedback on student outcomes (Carless, 2007; G. Gibbs & Simpson, 2004; Jessop et al., 2013; Kluger & DeNisi, 1996; Price, Handley, Millar, & O'Donovan, 2010; Pugh et al., 2017; Telio, Ajjawi, & Regehr, 2015; van der Ridder, McGaghie, Stokking, & ten Cate, 2015; Wu & Jessop, 2018), or students missing the low-stakes, learning opportunity nature of a formative assessment and perceiving it as high-stakes or consequential (Barrett et al., 2016; Bok et al., 2013; Broadbent, Panadero, & Boud, 2017; Heeneman et al., 2015; Massie & Ali, 2016; Pugh et al., 2017; Schut, Driessen, van Tartwijk, van der Vleuten, & Heeneman, 2018). In fact, numerous studies have shown assessment to negatively affect student learning behaviours (Cilliers et al., 2010; Cilliers, Schuwirth, Herman, et al., 2012; Cilliers, Schuwirth, et al., 2012a, 2012b; Harland, McLean, Wass, Miller, & Sim, 2014; Knight & Trowler, 2000).

The above literature gives evidence that assessment does indeed drive student learning, for it impacts what, when, how much and why students learn, which in turn influences what knowledge, skills, attitudes and competencies students develop and take with them into their profession. This has further consequences for patient care and public safety in the context of HPE. Yet, in order to use assessment to drive learning in desirable ways, the lecturer who is responsible for the practice of assessment needs to be considered. Yet, currently, lecturers' conceptions of assessment and what other factors impact on and shape their assessment practice currently remain largely unknown, especially in the global South.

3.3. Factors influencing lecturer assessment practice

HBT proposes that numerous and varied personal and contextual factors interact to influence behaviour. HPE literature was thus searched and the following factors as possible influencers of lecturer assessment practice were found (see **Table 3.3**.).

Table 3.3.: An overview of the personal and contextual factors influencing assessment practice.

Personal factors	Effect on assessment practice	References
Beliefs and conceptions	Varied	• Berendonk, Stalmeijer, and Schuwirth (2013) • de Jonge et al. (2017) • Halinen, Ruohoniemi, Katajavuori, and Virtanen (2013) • Harrison et al. (2017) • Postareff, Virtanen, Katajavuori, and Lindblom-Ylänne (2012) • Quesada-Serra, Rodríguez-Gómez, and Ibarra-Sáiz (2016) • Samuelowicz and Bain (2002)
Emotion and motivation	Barrier: negative emotions and low motivation Facilitator: positive emotions and high motivation	• Karthikeyan, O'Connor, and Hu (2019) • Kogan et al. (2017) • Massie and Ali (2016) • Myyry et al. (2019)
Assessment knowledge and understanding (literacy)	Barrier: low assessment literacy Facilitator: assessment literacy	• Deneen and Boud (2013) • Goos and Hughes (2010) • Medland (2018) • Norton et al. (2013)
Experience: number of years practicing assessment	Barrier: inexperienced Facilitator: experienced	• Goos and Hughes (2010) • Norton et al. (2013) • Quesada-Serra et al. (2016)
Professional identity	Varied	• Adendorff (2011) • Bartle and Thistlethwaite (2014) • Cantillon, Dornan, and de Grave (2019) • Stenfors-Hayes, Weurlander, Dahlgren, and Hult (2010)

Faculty or professional development and training, or an educational qualification	Facilitator	• Berendonk et al. (2013) • Karthikeyan et al. (2019) • Massie and Ali (2016) • Norton et al. (2013) • Popham (2009) • van Schalkwyk, Leibowitz, Herman, and Farmer (2015)
Contextual factor	**Effect on educational practice**	**References**
Organisational, institutional, departmental and disciplinary: values, knowledge, culture, traditions, policies and rules, leadership, peer-review, communities of practice	Barrier: unsupportive and resistant environment and colleagues Facilitator: supportive institutional culture and community of practice	• Bearman et al. (2017) • Deneen and Boud (2013) • Goos and Hughes (2010) • Kálmán, Tynjälä, and Skaniakos (2019) • Kogan et al. (2017) • Norton et al. (2013) • Skidmore, Hsu, and Fuller (2018) • van Lankveld, Schoonenboom, Volman, Croiset, and Beishuizen (2016) • van Schalkwyk et al. (2015)
Lack of academic recognition, regard (valuing) and reward (career progression) for educational practices – especially when compared to clinical work and patient care or research, which are then prioritised	Barrier	• Cantillon et al. (2019) • Huwendiek et al. (2010) • Norton et al. (2013) • van Lankveld et al. (2016) • van Schalkwyk et al. (2015)
Logistics and environmental constraints: lack of time, large class sizes, large workload, conflicting responsibilities, cost and availability of resources	Barrier	• Deneen and Boud (2013) • Goos and Hughes (2010) • Gilles, Detroz, and Blais (2011) • Kogan et al. (2017)

3.3.1. A number of personal factors influence lecturer assessment practice

Personal factors found in the literature include prestige or recognition (which is linked to motivation and value), gender, years of experience or expertise or qualifications (related to training and education, which in turn could be linked to perceived ability and confidence, or

agency and self-efficacy), which is related to knowledge of assessment (or assessment literacy), and conceptions of assessment as a potentially rich and important influencing personal factor.

In terms of prestige or recognition, lecturers who practice and teach at HEI hospitals found clinical work and biomedical research to be rated more highly and rewarded more often than teaching, causing faculty to focus their time and energy on those areas over education (Huwendiek et al., 2010; Knight & Trowler, 2000; Schiekirka-Schwake et al., 2017). Others reported a lack of recognition (Goos & Hughes, 2010) and a lack of incentives for lecturers to innovate assessment practices (Kogan et al., 2017; Medland, 2018; Norton et al., 2013).

Regarding gender and years of experience, Norton et al. (2013) reported that more experienced female lecturers were more likely to practice desirable, innovative assessment than males. The authors continued to say that lecturers who had completed a teaching qualification described a change in their views on assessment design and desired to improve their assessment practice, suggesting that exposure to educational training may be an effective tool to change assessment practice (Norton et al., 2013). This is supported by other studies; Goos and Hughes (2010) reported varying levels of confidence in subject coordinators (course convenors) regarding aspects of their assessment responsibilities, such as willingness to be innovative and try new approaches to assessment, and Huwendiek et al. (2010) described self-reported (perceived) expertise of medical educators as having general educational expertise but deficiencies in a number of areas – including assessment.

Similarly, Popham (2009) advocates teacher professional development programmes specifically focusing on improving teaching assessment literacy as currently the majority of teachers are "assessment illiterate" as they never received any assessment-focused educational training. Indeed, "assessment literacy" has more recently come to the fore in terms of educators lacking foundational assessment knowledge and skills, specifically recommending assessment education and training for educators to develop domain-specific knowledge, skills and assessment expertise over time (Bearman et al., 2016; Berendonk et al., 2013; Medland, 2018; Norton, Floyd, & Norton,

2019; Price et al., 2010; Quesada-Serra et al., 2016). Harden and Lilley (2018) recommend assessment training saying, “There is a need for a greater level of understanding and assessment literacy among all engaged with the education of healthcare professionals”. In fact, a scoping review reported that medical educators found faculty development programmes to be helpful in facilitating quality assessment practice (Karthikeyan et al., 2019).

Conceptions, as an example of a personal factor[4], warrants further attention as they have been shown to have a significant influence on behaviour (Arcila, 2018; Box, Skoog, & Dabbs, 2015; Meijer, Kuijpers, Boei, Vrieling, & Geijsel, 2016; Visser-Wijnveen, van Driel, van der Rijst, Verloop, & Visser, 2009). Moreover, there have been accounts that in order for a successful intervention in faculty training to occur, resulting in sustained, altered practice, core conceptions need to first be explicitly exposed, challenged and then transformed (Dawson et al., 2013; Harrison et al., 2017; Knight & Trowler, 2000; Meijer et al., 2016). Indeed, a number of authors describe how the success or failure of interventions to improve teaching and learning depended on whether or not educators’ conceptions of the purpose of assessment had changed. They suggest that fundamental to changing behaviour is first knowing and then changing those beliefs (Brown, 2004; Carless, 2007; Harrison et al., 2017; Ho, 2000; Ho, Watkins, & Kelly, 2001; Kane, Sandretto, & Heath, 2002; Kember, 1997; Pedrosa-de-Jesus & da Silva Lopes, 2011; Watkins et al., 2005).

3.3.1.1. Evidence that conceptions are an important personal factor in understanding behaviour and behaviour change

As HBT has its roots in (cognitive) psychology, which, on one, superficial level, explores how cognitive processes are determinants of behaviours, or, simply put, how thoughts can lead to action (later the sociological perspective was included to acknowledge the role of the environment or context) (Bandura, 1986; Eaton et al., 2003; Glanz et al., 2008). Several HBTs

[4] While “conceptions” does not appear as a specific personal factor in the most commonly employed HBTs, other, related terms (that the term conceptions includes and encompasses) are used; for example, “beliefs” is a prevalent factor in the Theory of Planned Behaviour (Munro et al., 2007).

include aspects of these mental influences (which I argue as “conceptions” in this thesis), for example, “beliefs” as a prevalent factor in the Theory of Planned Behaviour (Munro et al., 2007).

In investigating the literature, numerous terms have been used interchangeably to describe conceptions. These include perceptions (de Jonge et al., 2017; Knight & Trowler, 2000; Quesada-Serra et al., 2016; Trigwell, Prosser, & Waterhouse, 1999), beliefs or belief systems (Brown, 2004; de Jonge et al., 2017; Halinen et al., 2013; Postareff et al., 2012; Samuelowicz & Bain, 2002), orientations (Bolander, Josephson, Mann, & Lonka, 2006; Halinen et al., 2013; Postareff et al., 2012; Samuelowicz & Bain, 2002), perspectives (Berendonk et al., 2013; de Jonge et al., 2017), views (Bolander et al., 2006; Halinen et al., 2013), ways of seeing and understanding (Halinen et al., 2013), interpretations, thoughts, mental structures (Brown, 2004), personal theories or philosophies (Trigwell et al., 1999), and metaphors (Rees, Knight, & Cleland, 2009). Indeed, the lack of a definition for conceptions in these papers underscores the need to define the term. In this research project, the term “conception” is aligned with the Phenomenographic definition for the following reasons described below.

Conceptions refers to the qualitatively different ways in which individuals experience, understand and make meaning of an aspect of reality or a phenomenon, or the varied interactions and interpretations between a subject (person) and an object (phenomenon) in the real world, or, put simply, descriptions of lived experiences (Ashworth & Lucas, 1998; Entwistle, 1997a; Marton, 1981). One author defines it well, “‘Conceptions’ is the term used… to describe the organizing framework by which an individual understands, responds to, and interacts with a phenomenon. The structure of teachers' conceptions is not uniform and simple; they appear to be multifaceted and interconnected” (Brown, 2004). It may then be helpful to picture conceptions as multiple and complex nodes that are connected and interlinked (see **Appendix 3.3.**). To this end, literature referring to conceptions, and its many related terms, was carefully analysed to determine if, in fact, the authors were speaking of conceptions, as defined above, and only included if synonymous.

In terms of conceptions impacting on behaviour, studies have shown that how a teacher thinks about teaching and learning influences how they teach, how their students learn and the learning outcomes their students achieve. Indeed, "Teachers' conceptions of teaching have a strong impact on approaches and practices, and because of this teachers do not adopt approaches to teaching that reach beyond the sophistication of their conceptions" (Kálmán et al., 2019). For instance, there is evidence that if a teacher adopts an information-transmission and teacher-centred approach to teaching, students adopt a surface-approach to learning, whereas, if teachers adopt a student-centred approach to teaching, students adopt deep-learning strategies (Entwistle & Peterson, 2004; Hartman & Nelson, 1992; Ho et al., 2001; Kember, 1997; Ross, 2017; Samuelowicz & Bain, 2002; Trigwell et al., 1999; Watkins et al., 2005). There have also been reports of students' conceptions influencing their learning behaviour (Al Kadri, Al-Moamary, & van der Vleuten, 2009; Cilliers et al., 2010). As Trigwell et al. (1999) found educators' teaching beliefs and practices to be linked, so it is reasonable to postulate that this may be true for assessment as well.

There are indications that assessment conceptions and behaviours are related in both school (Basic Education) (Brown, 2004; Chen & Bonner, 2016; DeLuca, Coombs, & LaPointe-McEwan, 2019) and university (HE) level settings (Cilliers et al., 2015; Cilliers & Tekian, 2016; Halinen et al., 2013; Harrison et al., 2017; Huwendiek et al., 2010; Postareff et al., 2012). Postareff et al. (2012), Halinen et al. (2013) and Brown (2004) contend that understanding teachers' conceptions of assessment are vitally importantly because it reflects their beliefs on all pedagogical activities, such as teaching, learning and curriculum design.

A school-level study by Brown (2004) explored what conceptions teachers held regarding the purpose of assessment through a questionnaire. Teachers agreed that assessment is used to improve teaching and learning, for school accountability, and to a lesser extent, for student accountability (Brown, 2004). No teachers viewed assessment as irrelevant (Brown, 2004). A university-level study investigated how lecturers and learners perceived workplace-based assessments and it was found that they held various conceptions of assessment, including

"agency" (assessment should be used to guide and direct learning through formative assessment and feedback), "mutuality" (assessment should be embedded in the learning process and both lecturers and learners are responsible for this learning process through constructive collaboration and longitudinal mutual relationships), "objectivity" (assessment is both formative and summative, serving an auditing role, assessors should be experienced and assessment should have clear criteria), "adaptivity" (there is flexibility in terms of the assessor and assessment practice) and "accountability" (assessment serves as a form of quality control) (de Jonge et al., 2017). The study concluded that awareness of these various conceptions of assessment are important for practice as the gap between actual and ideal practice may be better understood and future professional development interventions designed as the acceptance, use and effectiveness of different assessment programmes may depend on lecturers' conceptions of assessment (de Jonge et al., 2017).

In a HEI setting, Halinen et al. (2013), who states that assessment is an integral part of the teaching process as it is able to substantially influence student learning, investigated how lecturers viewed assessment in relation to their teaching practice, how they justified their assessment-related decisions and how they thought of their and their students' roles in assessment processes. Through reflection and interviews the authors found that in the areas of assessment practice, student learning and the purpose of assessment, lecturers held a range of conceptions: "status quo", "awareness" and "development" (Halinen et al., 2013). The first category ("status quo") refers to a lack of pedagogical awareness and rather a focus on content, and an unthinking adoption of conventional or traditional practices without questioning the validity or suitability of the assessment method regarding what was to be measured (Halinen et al., 2013). Lecturers with this conception of learning viewed themselves as autonomous, the guardians of knowledge, and thus interacted little with learners, using assessment to measure whether or not students had assimilated their teaching or not (Halinen et al., 2013). The second category ("awareness") refers to being conscious of assessment, its influence and the various role players, focusing on both content and the learner (Halinen et al., 2013). Importantly, lecturers with "awareness" conceptions of assessment are able to critically engage with assessment and

its practice, yet are unable to take action and implement change themselves (Halinen et al., 2013). This action takes place in the third category ("development") in which lecturers are motivated to move beyond awareness to action, focusing on the learner, developing a non-hierarchical relationship with them ("students as partners"), encouraging critical-thinking and input from learners, using assessment to monitor learner-progress and adapt teaching processes on an ongoing basis, continually working to improve future assessment methods to foster lifelong learning. The goal of these lecturers is to use assessment creatively, empower students to apply knowledge and learn alongside assessment practices (Halinen et al., 2013). The authors continue to suggest that pedagogical education (related to assessment training and increasing assessment literacy), peer and institutional support, as well as interactions within the HE environment may play an important role in influencing assessment conceptions and practice (Halinen et al., 2013).

Postareff et al. (2012) provides evidence of the link between conceptions of assessment and assessment practice in a HEI study. Faculty with conceptions of transmission-teaching[5] practice traditional assessment methods (measuring repetition, memorisation of facts and content coverage); whereas, faculty with constructivist conceptions practice more continuous, alternative assessment methods (measuring application of knowledge, deep understanding and the process and development of students' own thinking) (Postareff et al., 2012). A related finding was found in terms of assessment grading practice: school teachers' constructivist conceptions of teaching and learning led them to an "academic enabling" form of grading. Teachers with constructivist conceptions graded assessments in such a manner as to guide and motivate students to learn, taking personal knowledge and understanding of the students and the context into consideration (Chen & Bonner, 2016).

[5] See review by Kember (1997) on teachers' conceptions of teaching and learning (mentioned in **Appendix 3.3.**). These conceptions are described along the continuum of information-transmission/teacher-centred to apprenticeship/student-teacher interaction to conceptual change/student-centred. At one end of the spectrum teachers play the lead role in passing structured information on to their students (transmission-teaching), moving towards facilitating understanding and intellectual development at the other end, where the student is the responsible party for learning (constructivism) (Kember, 1997).

While conceptions have been shown to have an impact on lecturer assessment behaviour, so other studies have also shown that conceptions alone are not enough to drive assessment behaviour as other factors also shape practice (Bearman et al., 2016; Deneen & Boud, 2013; Harrison et al., 2017; Murray & Macdonald, 1997; Offerdahl & Tomanek, 2011). Indeed, authors suggest that both conceptual change and a deep dissatisfaction with or change to the institutional or departmental culture and practice, with educator involvement and buy-in, needs to accompany assessment changes in order for a lasting and genuine changes in assessment practice to occur (Bearman et al., 2016; Deneen & Boud, 2013; Offerdahl & Tomanek, 2011).

Yet, before the impact of contextual factors is discussed, an additional personal factor that emerged during data collection and analysis was that of identity. Indeed self-identity has been found to be an important factor in HBT Theory of Planned Behaviour in particular (Chatzisarantis, Hagger, Wang, & Thøgersen-Ntoumani, 2009; Fielding, McDonald, & Louis, 2008; Hagger & Chatzisarantis, 2006; Terry, Hogg, & White, 2010). Identity-formation and development also involves interaction between cognitive (individualist) and social (relational) processes (Beauchamp & Thomas, 2009; Cantillon, D'Eath, de Grave, & Dornan, 2016; Cantillon et al., 2019; Monrouxe, 2010), similar to HBT's internal (personal) and external (contextual) factors.

Since identity-formation and development has a cognitive component (the social processes are mentioned below, but also overlap with the contextual factors discussed in **Section 3.3.2.**), in this study I argue that identity is an important element of a lecturers' *conceptions* of assessment; it relates to how a lecturer views themselves (identity as a cognitive organising frame or analytical lens) and their roles *in relation to assessment* (Beauchamp & Thomas, 2009). So to specifically explore their assessor identity and role (or lack thereof), will assist us in making sense of their broader conceptions of assessment.

3.3.1.2. Lecturer identity is a personal factor of significance as an element of conceptions

In searching the literature for lecturer professional development, literature in HPE often speaks of medical students developing their professional identities as clinicians or medical practitioners, and the role clinician educators (medical teachers) play in this process, yet less literature exists in describing the development of an educator or clinician-educator identity, and especially that of an assessor.

Professional identity refers to the way in which individuals understand and view themselves, and their experiences, and how they wish to present themselves to others (Lieff et al., 2012; Rosenblum, Kluijtmans, & ten Cate, 2016; Steinert, O'Sullivan, & Irby, 2019). Identity has been defined as, "The values, beliefs, sense of affiliation, aspirations, and synchrony with the norms of… (said) profession" (Rosenblum et al., 2016), for instance, belonging to an "academic tribe" (Knight & Trowler, 2000). The general acknowledgement is that identity is both a product and process (dynamic and adaptive), multi-faceted, consisting of individual cognitive or psychological and social relational or collective aspects, thereby situating a person in context (self in relation to society or others) (Beauchamp & Thomas, 2009; Cantillon et al., 2019; R. Cruess, Cruess, Boudreau, Snell, & Steinert, 2014; Rosenblum et al., 2016; Steinert et al., 2019).

Professional academic identity formation may involve both individual and collective (social) processes (van Lankveld et al., 2016). Individual processes include reflective practice and the use of personal-narratives (Wald, 2015). Social processes include socially-situated, collaborative and experiential learning (R. Cruess et al., 2014), communities of practice (Adendorff, 2011; Andrew, Ferguson, Wilkie, Corcoran, & Simpson, 2009; Cantillon et al., 2016) and important relationships, such as role-models and mentors (S. Cruess, Cruess, & Steinert, 2008; Rosenblum et al., 2016; Wald, 2015).

In general, it has been proposed that identity formation is similar to social and transformative learning approaches whereby individuals are exposed to a new idea; there is a disruption, disorientation, tension or turbulence, followed by acceptance, adoption and assimilation of new knowledge, skills, attitudes, values and behaviours (Mezirow, 1990). This awareness and subsequent change has also been related to critical reflection, for it supports transformative learning and can lead to changed beliefs, behaviours and identity (Meijer et al., 2016). For example, reflection allows for qualitatively shifts in the way in which individuals see themselves and their work (Helsing, Howell, Kegan, & Lahey, 2008); it helps them to make sense of cognitive disequilibrium triggered by new experiences (Kay, Berry, & Coles, 2019); and facilitates the construction of their personal theories of education (Yan Fung, 2005).

One study simplified both individual and social identity formation processes to a simple two-step process of external and internal change, external change related to stimulation by a new idea or situation (such as a recent appointment as an educator) followed by an internal change of becoming and being (experiencing a personal calling and development into the new profession) (Stone et al., 2002). Indeed, many clinicians "become" clinician-educators "on the job" or through experiential learning, developing craft knowledge and implicit expertise, as most clinicians do not possess formal educational training, as has been previously mentioned (Cantillon et al., 2019; R. Cruess et al., 2014; Swanwick, 2008).

Another suggested mechanism was based on models of tolerance and management of ambiguity and complexity. Dual identity formation could take the form of intersection (an individual possesses a number of differing and distinct identities that intersect with each other), dominance (an individual expresses a primary identity with secondary identities), compartmentalization (an individual expresses different identities based on the context or situation they are in), or merger (whereby an individual combines and expresses a sum total identity) (Rosenblum et al., 2016). In HPE specifically, there have been reports of teachers have varied understandings of their professional identity and development. Some saw their identities and roles as educators and clinicians as separate, whereas others saw these as one (Stenfors-Hayes, Hult, & Dahlgren, 2012),

while others viewed their educator identity as secondary to their clinician identity It was suggested that, as multiple professional identities develop, they move along a continuum from compartmentalised to merged (Steinert et al., 2019). In fact, a recent review reported that clinicians were able to reconcile their educator identities through first juggling these conflicting identities, finding mutuality between them, before integrating the identities into a merged clinician-educator identity (Cantillon et al., 2019).

While the exact processes or mechanisms of identity development remain open to further investigation, a number of influencing factors, both enabling and constraining of clinician-educator identity development, have been identified in a number of studies in various ways: personal factors, relational factors, experiential factors, contextual factors and career interests (Lieff et al., 2012).

A number of enabling factors have been reported in the literature. An educator identity is able to develop and strengthen in individuals with aligned values, a sense of appreciation, a felt responsibility and commitment to education (van Lankveld et al., 2016; Wald, 2015), as well as a motivation to provide better education for future colleagues and improve patient care (Bartle & Thistlethwaite, 2014; Blitz, Bezuidenhout, Conradie, de Villiers, & van Schalkwyk, 2014; Stone et al., 2002). Personal resilience, in terms of tolerating and managing ambiguity, was found to increase as an individual's professional identity evolved from simple to more complex (Rosenblum et al., 2016; Wald, 2015). This is related to a self-perceived competence and confidence in the emerging identity which encourages further development (Beauchamp & Thomas, 2009; Rosenblum et al., 2016; van Lankveld et al., 2016). This growth in confidence and capacity could be due to a familiarity with or an increase in education-related knowledge and skills (Åkerlind, 2003; Stone et al., 2002). Providing educational opportunities, such as training programmes, for individuals to learn has been suggested as important for identity development (Helsing et al., 2008; Rosenblum et al., 2016; van Schalkwyk, Cilliers, Adendorff, Cattell, & Herman, 2013).

Identity development is helped by supportive relationships (peers, role-models, mentors, supervisors) and networks (communities of practice), at departmental, disciplinary and HEI levels, as they provide a sense of belonging and connectedness with a wider educational community (Bartle & Thistlethwaite, 2014; Lord et al., 2012; Rosenblum et al., 2016; van Lankveld et al., 2016). Similarly, if the working environment is empowering, then individuals feel supported in their journey. Furthermore, if teaching was valued, or there was an associated recognition and reward, individuals' educator identities were strengthened (van Lankveld et al., 2016).

Several constraining factors have also been identified. If an increase in knowledge and skills, and associated with a sense of agency, encourages educator identity development, then, conversely, when individuals are unfamiliar with the discourse of a new field and insecure in the role, for example lack the "language" to participate, this hampers their identity formation (Adendorff, 2011; Blitz et al., 2014; Stenfors-Hayes et al., 2010). Similarly, if a department or HEI does not value education, which leads to a lack of recognition and reward of educational practice, a low career status or promotion, then individuals are discouraged from pursuing it further (Adendorff, 2011; Bartle & Thistlethwaite, 2014; Stenfors-Hayes et al., 2010). This relates to professional hierarchy and community of practice legitimisation, and a lack thereof for educators, as traditional disciplinary research or clinical practice is valued and privileged over teaching, which can lead individuals to adapt their educator identity to suit institutional norms and departmental culture (Cantillon et al., 2016; Cantillon et al., 2019; van Schalkwyk et al., 2015).

Lecturers have expressed experiencing mixed and conflicting messages when it comes to the role and value of clinician-educator. Teaching is critically important yet research is prioritised and privileged in terms of recognition and career promotion (Adendorff, 2011; Cantillon et al., 2019; van Schalkwyk et al., 2013; van Schalkwyk et al., 2015). Lecturers in several studies expressed these sentiments: "The success of clinical educators is measured in terms of research productivity and clinical service rather than teaching" and described feelings of, "Inauthenticity and marginalisation for those clinicians who favour teaching over research" or clinical service (Bartle

& Thistlethwaite, 2014). One study put it this way: “She (a clinician-educator) feared that colleagues would disregard her work in the area of teaching as “unscientific”. She also expressed concern that her Head of Department’s position of support could change if too many people in their department started doing classroom research (as an “easy option” alternative to disciplinary research), especially if those projects lacked scientific rigour” (van Schalkwyk et al., 2013). Moreover, “How emerging scholars are positioned with respect to disciplinary research could thus influence the origin of identity issues. Those who are active in the scholarship of teaching and learning (SoTL) and disciplinary research, it seems, can still protect their status through their disciplinary research. Similarly, emerging scholars remaining in their disciplines express less fear of losing their disciplinary identity” (Adendorff, 2011). Thus, this tension remains; the disconnect or competition between two professional academic identities, the clinician or disciplinary researcher and the educator.

Part of this tension, mentioned briefly above, is the idea of credibility and rigour, and the perceived risk to their careers, which relates to numerous HBTs that have element of risk and perceived benefits and barriers, such as The Health Belief model (Cilliers et al., 2015; Munro et al., 2007). For example, when interviewing clinicians who were transitioning over into performing a more educator role, the clinicians still prioritised their identities as clinicians over educators: “(The medical educators) emphasised the importance of *continuing* with their clinical identity in terms of *credibility* as an educator and their commitment to patient care” for they expressed either that their identity and role as a clinician bestowed authenticity, relevance and authority on their teaching or that they were primarily clinicians with an “interest” in education, or “(They) wished to be seen as doctors with an interest and skill in education while working as specialists” (Bartle & Thistlethwaite, 2014). Similarly, “Participants (clinician educators) were concerned that their increasing commitment to their academic role might compromise their clinical competence” (Lieff et al., 2012).

This is in line with a systematic review on HEI teacher identity formation where professionals who had transitioned into academia, still strongly identified with their profession and not their newer

role as an HEI educator, especially during the early years of their HEI careers, they, "(Considered) professional expertise important to their credibility as teachers" (van Lankveld et al., 2016). This too was observed in a review on teacher identity in the medical profession; reputation and legitimacy depended on the clinician identity and practice (Cantillon et al., 2019). Relatedly, clinician-scientists experienced feelings of inferiority in their dual-identity as they perceived themselves to be, "Less expert in either clinical medicine or research than peers focused solely on either domain." Thus faculty programmes designed to highlight the unique value of possessing expertise on more than one discipline, community-building (peer-support) and role modelling and mentorship were found to support dual identity development (Rosenblum et al., 2016).

Just as conceptions alone cannot account for lecturer assessment practice, so, while identity emerged as an important personal factor and element of lecturers' conceptions of assessment, it too is influenced by additional social processes and contextual factors. As such, contextual factors impacting on lecturer assessment practice will now be detailed.

3.3.2. Contextual factors influencing lecturer assessment practice

As HBTs expanded from its psychological roots to include insights from sociology, factors beyond the personal and conceptual were included in models, for an individual does not exist within a vacuum but is subject to contextual influences.

Several proximal contextual factors have been described including workload or a perceived lack of time to dedicate to assessment practice, which includes factors that lead to workload pressure, such as increasing student numbers or large class sizes; perceived cost, access to resources and facilities (Deneen & Boud, 2013; Gilles et al., 2011; Goos & Hughes, 2010; Kogan et al., 2017); collegial and managerial support, disciplinary and organisational or institutional cultures and policies (Bearman et al., 2017; Norton et al., 2013; Skidmore et al., 2018).

Pragmatic issues relating to workload, such as the time needed to mark assessment, can have a profound impact on the focus of assessment: “Measurement and accreditation of learning generally takes priority and the intense focus on marking has led to a reduced emphasis on assessment as a vehicle for learning” (Price, Carroll, O’Donovan, & Rust, 2011). In terms of class size, in an international survey conducted in Belgium and Canada investigating assessment practices, found that the larger the class size the greater the constraints. In smaller class sizes longer-answer question examinations were selected more often than in larger class sizes in which short-answer question examinations, such as MCQs, were favoured (Gilles et al., 2011). In other instances, lecturers have resisted changing assessment practice due to a fear of increasing workload (Norton et al., 2013).

There are also reports that assessment behaviour is influenced by the existing organisational, institutional culture and peer practice of assessment; practicing the “status quo” or what has always been done (Bearman et al., 2016; Halinen et al., 2013; Harrison et al., 2017; Jessop & Tomas, 2016; Johnson, Scholes, & Whittington, 2005; Kogan et al., 2017; Norton et al., 2013; Segers & Tillema, 2011). Indeed, institutional culture was reported to influence assessment practices by applying resistance towards lecturers and their desire to change assessment practices (Bearman et al., 2016; Boud, 2000; Deneen & Boud, 2013; Harrison et al., 2017; Kogan et al., 2017; Offerdahl & Tomanek, 2011). Others have described in detail different institutional cultures of assessment: a culture of fear (where innovation is stifled or punished and the norm maintained); a culture of compliance (where accountability and accreditation are the priority and there is little commitment to student learning); a culture of evolving student learning (where educators are transitioning towards improving student learning), and a culture of student learning (Skidmore et al., 2018).

One study found that when faculty worked in groups to develop consensus on a new assessment method, they were more confident, comfortable and likely to use it, expressing that they felt validated and a part of a community (Kogan et al., 2017). Moreover, others have described the important role of the departmental, disciplinary or working faculty group contexts as being

particularly powerful ecosystems that are conducive for change as individuals are able to participate in an educational community of practice. Knowledge, skills and attitudes may be developed, a healthy culture and supportive and enabling environment are created, and an increase in prestige and reward are felt (Kálmán et al., 2019; Knight & Trowler, 2000; Mälkki & Lindblom-Ylänne, 2012; Medland, 2018; Prebble et al., 2004; Schiekirka-Schwake et al., 2017; Skidmore et al., 2018; Witman & Richlin, 2007). In describing how various departmental cultures can lead to faculty change individualistic departments make lecturers feel isolated and unwilling to collaborate (departments consist of separate and competitive cliques) and departments with a culture of contrived collegiality make faculty collaboration feel more formal and regulated, whereas, in departments of professional learning communities, collaboration and change are voluntary and more flexible (Kálmán et al., 2019).

An example of this in an African HEI setting (Rwanda) is students who expressed a desire for assessment questions that tested their deep thinking and analysis skills. However, due to the pressure lecturers faced from managers to pass students to keep their failure rates low, lecturers practiced an assessment that and tested students' memorisation of knowledge skills over critical thinking (Niyibizi et al., 2018). This and other studies suggest that in order to change assessment practice institutional and departmental culture and resistance need to be overcome (Deneen & Boud, 2013; Harrison et al., 2017; Kogan et al., 2017). Buy-in and support are needed, followed by convincing key stakeholders of the need for change, collegial collaboration and consensus, implementation of supportive institutional policy, and facilitation of scholarship and educator training and development (Harrison et al., 2017; Medland, 2018; Offerdahl & Tomanek, 2011).

Additionally, subject discipline and institution type were found to be contextual factors that influence assessment practices. So-called hard disciplines (engineering and science) and more traditional (established and older) institutions had a negative impact on lecturers' assessment practices, contributing to constraining innovation and implementation of desirable assessment practices compared, to so-called soft disciplines (humanities and social science) and more modern institutions (Norton et al., 2013). Similarly, another study found that in soft-disciplines,

lecturers were more open to professional development than those in hard-disciplines who resisted proposed change interventions (Kálmán et al., 2019). Yet, interestingly, other studies showed that the assessment differences were greater between HEIs than discipline or subject, suggesting that the management of the assessment regime may have a greater impact on student learning than what they are learning (G. Gibbs & Dunbar-Goddet, 2009; Price et al., 2011).

Distal contextual factors include structural factors and national culture. These structural factors include the social, economic, political and legal environment of a place. In this study, we have summarised said structural factors under the term of "resource-rich" or" developed" and "resource-constrained" or "developing" contexts (which will be discussed in **Section 4.2.** below). Considering structural factors, workload may be an influencing factor yet, in a resource-rich context, workload could also be influenced by the factor of clinician-educators needing to see a certain number of patients in order to bring in a threshold amount of revenue at the university-hospital in order for it to remain open. In a resource-constrained context, workload could be influenced by the disease burden and reality of patients who only have access to public university-/hospitals (Mullan et al., 2011). Even though the doctor:patient ratio (dr:pt) is lower in developing countries than developed, there is further maldistribution of clinicians between urban (higher ratio) and rural (lower ratio) areas, compounded by an additional skewed distribution between private (higher ratio) and public (lower ratio) hospitals in mostly urban areas (Human Resources for Health South Africa, 2011).

Yet, in light of using Southern Theory as a conceptual framework, the distal context of the global South extends beyond that of pragmatic or structural descriptions of gross domestic product (GDP) and disease burden, and looks at context through the lens of power. For example, the underlying history of colonialty and lasting legacy of inequality and discrimination have shaped national culture.

National culture, the shared values, beliefs, norms, traditions of a nation, unsurprisingly varies from one context to the next. One would then expect differing impacts of national culture on

assessment practice from one setting to another. Studies have shown, using Hofstede's model of national culture[6], that cultural dimensions have had different effects on assessment practice in different countries (Earley & Stubblebine, 1989; McLeay & Wesson, 2014; Ng, Koh, Ang, Kennedy, & Chan, 2011; Varela & Premeaux, 2008).

In an HPE setting, researchers compared perceived feedback instructiveness during clerkship between Indonesian and Dutch student populations. Indonesian students perceived feedback to be more instructive when it was jointly initiated by the supervisor and student and came from a specialist whereas Dutch students perceived feedback as more instructive when it was based on observation (Suhoyo et al., 2017; Suhoyo, van Hell, Prihatiningsih, Kuks, & Cohen-Schotanus, 2014). The authors explain the results of this study by using Hofstede's model of the six dimensions of national culture (see **Appendix 4.1.** for more details on Hofstede's model of national culture) (Hofstede, 2001, 2011; Hofstede & McCrae, 2004; Hofstede & Minkov, 2010). For example, Indonesia has a high Power Distance score compared to the Netherlands, implying

[6] Hofstede's model of (static) national and organisational culture describes the culture of a country according to six dimensions: power distance, individualism, masculinity, uncertainty avoidance, long-term orientation and indulgence. Power distance refers to the power distribution (high score accepts hierarchy, low scores desires equality); individualism refers to the strength to which an individual has ties to a community (high score favours individualism, low score favours community); masculinity refers to how society views patriarchally-dominated gendered roles (high score favours patriarchally-dominated gendered roles, low score favours equal gendered roles); uncertainty avoidance refers to the degree to which ambiguity or change are tolerated (high score indicates rigid codes of belief and behaviour, a low score indicates a more relaxed attitude to deviations); long-term orientation refers to how society values long-term traditions and values (high score favours tradition, low scores favours innovation); and indulgence refers to the extent to which desires are controlled (high score indicates self-control and restraint, low score indicates gratification of desires) (Hofstede, 2001, 2011; Hofstede et al., 2010). A similar model of culture, the GLOBE study, is mentioned in **Appendix 4.1.** There have been critiques of Hofstede's model: that it's measurements are based on static scores (culture may be dynamic) and inferences from individuals scores to a collective level cannot be made (Javidan, House, Dorfman, Hanges, & de Luque, 2006; Signorini, Wiesemes, & Murphy, 2009; Tung & Verbeke, 2010). However, in this study, Hofstede's model of national culture is only being consulted for data analysis and commentary purposes.

that power hierarchies exist and are maintained thereby explaining why Indonesian students would find feedback from specialists (who are high in the hierarchy and thus possess authority) as more instructive, versus the Dutch students, with a low Power Distance score (favouring equality over hierarchies), who found feedback from specialists and residents as being equally beneficial (Suhoyo et al., 2017; Suhoyo et al., 2014).

Additionally, when a Western (Canadian) feedback assessment tool (in-training evaluation reports) was adopted in a non-Western context (Qatar), students were reluctant to provide constructive critique towards their supervisors due to the perceived power relations, in accordance with Hofstede's measures of Power Distance and Collectivism, indicating the difference in outcomes of applying the same assessment method in different cultural settings (Wilbur, Bacha, & Abdelaziz, 2017).

To summarize, as assessment in HPE drives learning, with significant consequences extending beyond the student to the patient, it is the desire of educators to use assessment to drive outcomes in a beneficial manner. However, little is known how the lecturers impacts on this process. Lecturer assessment practice is the focus of this study, for, I argue that lecturer assessment practice is a behaviour that needs to first be understood (how they practice assessment; what and why?), before it may potentially be targeted and changed through evidence-based and theory-informed faculty training and continued professional development interventions. The need for changed assessment preparedness is highlighted by reports on the need for assessment education for underprepared and poorly practicing assessors (Alexander et al., 2012; Burchard et al., 1995; Daelmans et al., 2005; Holmboe, 2004; Huwendiek et al., 2010; Kogan et al., 2017; McGill et al., 2013; Medland, 2018; Popham, 2009; Schiekirka-Schwake et al., 2017).

The HBT conceptual framework employed in this study suggests that numerous and varied personal and contextual factors interact to influence lecturer assessment behaviour, such as, prestige or recognition, gender, years of experience or expertise, qualifications and assessment

literacy, perceived ability and confidence, workload pressures, environmental or structural barriers, and culture (disciplinary, institutional and national). Conceptions, as a particular personal factor of interest, appears to be important for understanding and potentially changing behaviour (Brown, 2004; Carless, 2007; Harrison et al., 2017; Ho, 2000; Ho et al., 2001; Kane et al., 2002; Kember, 1997; Pedrosa-de-Jesus & da Silva Lopes, 2011; Watkins et al., 2005). Professional identity, I argue as an element of conceptions, also emerges as a key influencing factor on lecturer assessment practice.

The literature has also clearly shown that no single factor, such as conceptions or identity alone, may describe or explain lecturer assessment practice (Bearman et al., 2016; Deneen & Boud, 2013; Harrison et al., 2017; Mälkki & Lindblom-Ylänne, 2012; Murray & Macdonald, 1997; Offerdahl & Tomanek, 2011). For instance, several examples show a mismatch between lecturers' conceptions of assessment and their enacted practice due to the influence of additional personal and contextual factors (Box et al., 2015; Brown, 2004; Entwistle & Peterson, 2004; Kane et al., 2002; Mälkki & Lindblom-Ylänne, 2012; Murray & Macdonald, 1997; Pedrosa-de-Jesus & da Silva Lopes, 2011; Pereira, 2016). Indeed, assessment practice is stated to be the result of "(The) complex social nature of interwoven personal and environmental influences… within the constraints and affordances of a local environment" (Bearman et al., 2017; Henderson, Ryan, & Phillips, 2019). One study described this, "Complex interaction between teachers' beliefs, knowledge, and practices" as an interplay between teachers' conceptions of assessment, their expectations, habits, pressures felt, motivations, student dispositions and contextual elements (Box et al., 2015).

Lecturers may have formative- and learner-centred conceptions of assessment, yet practice a more teacher-centred, traditional summative assessment due to other influencing factors such as, "(A) lack of material and human resources, the teacher:student ratio, heavy workload, lack of availability to the assessment process, time spent for doing research and institutional constraints" (Pereira, 2016). Indeed, in a study conducted at 18 Spanish HEIs, from all disciplines, a disconnect between lecturers conceptions of assessment and their assessment practice was

found. Lecturers viewed assessment as important for student learning, yet failed to practice formative assessment (Quesada-Serra et al., 2016). Several reasons for this disconnect were given, including a lack of assessment literacy and assessment training, as well as restrictions imposed by institutional bureaucracy in the form of HEI rules regarding assessment (Quesada-Serra et al., 2016). Similarly, a study in the Netherlands found that lecturers' conceptions of assessment and their outworking in their assessment practice was moderated by their level of assessment knowledge and abilities, their assessor identity (ownership of the assessment task), colleagues, institutional assessment rules and assessment context (Berendonk et al., 2013).

Thus, while HBT provides a helpful and valid framework within which to study lecturer assessment behaviour, the second theoretical perspective consulted for this project, Southern Theory, needs to be considered, especially as the majority of literature reviewed in this chapter originated from the global North, which provides a poor representation of how the majority of the world's population lives (Connell, 2014; Connell et al., 2018a, 2018b; Gosselin et al., 2016; Greysen et al., 2011; Henrich et al., 2010; Mullan et al., 2011; Rotgans, 2011; Tutarel, 2002; Walubo et al., 2003).

While HBT is a "Northern Theory", it does consider placing behaviour in context and context, as observed with national culture, can have effects on how assessment is perceived and practiced. Indeed, a number of studies have stated how essential it is for assessment to be contextually-appropriate, being mindful of pragmatic concerns, such as structural barriers (resource-constraints) prevalent the global South (Gosselin et al., 2016; Tsai et al., 2012; Walubo et al., 2003).

Beyond the practical concerns lies deeper theoretical issues. If assessment is a socially-situated practice with the existence of power relations, then adopting Western or Northern assessment theory and practice in the global South could lead to social injustice and oppression (Hanesworth et al., 2018; Kester, 2018; Luckett, 2016; Maldonado-Torres, 2016; Ndlovu-Gatsheni, 2013; Shay, 2004; Shay & Peseta, 2016). There is a need to develop evidence-based and theory-informed

faculty training programmes, for there is evidence that these validated models may result in more effective training transfer and behaviour change (Bahar-Ozvaris et al., 2004; Grossman & Salas, 2011). However, where that theory is produced, by whom and for who, remains of critical importance, especially as empowerment of Southern lecturers to practice assessment powerfully is the goal (Connell, 2014; Connell et al., 2017; Connell et al., 2018a, 2018b; Fawcett & Hearn, 2004; Ndlovu-Gatsheni, 2013; Shay & Peseta, 2016).

3.4. Problem Statement, Aims and Objectives

Assessment drives learning. Assessment thus has consequences for the student, with often unintended negative outcomes, and in the context of HPE, the public. The desire is for assessment to be used in such a way as to contribute to the creation of powerful (and socially just) learning environments, so that assessment opportunities are learning opportunities for the student. Critically, this may not take place without the individual who practices assessment – the lecturer. Yet, at present, little is known about lecturer assessment practice and how this impacts on this assessment process, especially in the global South. Furthermore, there have also been reports of poor assessment literacy and a need for validated, evidence-based faculty training.

This research study proposes that in order to understand how lecturers impact assessment and how they may contribute to its powerful practice, their assessment practice (what and why) needs to be understood. To this end, I propose that lecturer assessment is a behaviour that may be described (and changed, in the future). I believe HBT to be a helpful conceptual framework to study this behaviour, as literature has shown numerous and varied personal and contextual factors interacting to impact on enacted lecturer assessment practice. Southern Theory was also consulted as a theoretical perspective in this study as it moves beyond mere description of lecturer assessment practice to consider the deeper issues of knowledge creation and social justice in assessment, especially as the vast majority of theory and practice of assessment originates from the global North.

The eventual outworking of this study would be the development of a novel, evidence-based HBT describing and explaining lecturer assessment practice in the global South, which, in turn, may then inform potential faculty training and continued professional development interventions to specifically target valid and relevant factors, and, ideally, effectively lead to a powerful (and socially just) Southern lecturers and lecturer assessment practice with enhanced student learning and public safety. The findings of this research project too would contribute to Southern Theory, providing an alternative knowledge to the dominant discourse.

3.4.1. Problem Statement: HPE lecturer assessment practice in the global South

Before lecturer assessment practice may be changed, it first needs to be understood. To this end, the aim of this research study is to explore lecturer assessment behaviour is diverse Southern contexts, particularly looking at what personal and contextual factors impact on their practice, in accordance with HBT, with a special focus on lecturers' conceptions of assessment as a potentially important element.

3.4.2. Aims and Objectives

The objectives of this research are:

Objective 1

To explore lecturers' conceptions of assessment in HPE programmes in a diverse range of Southern contexts.

Objective 2

To explore what factors, additional personal and contextual, influence lecturers' assessment practice in HPE programmes in a diverse range of Southern contexts.

Chapter 4: Methodology, Methods and Analysis

4.1. Methodology (Research Design)

4.1.1. Research Paradigm and approach

An interpretivist paradigm and a qualitative research approach were used in this study. Within an interpretivist research paradigm reality is socially negotiated and knowledge is co-constructed between individuals and society (Stalmeijer, McNaughton, & van Mook, 2014). One of the aims of interpretivist paradigm is to understand phenomena, specifically how and why people experience the world around them in terms of thoughts, feelings and opinions (Bunniss & Kelly, 2010). Using an interpretivist paradigm and qualitative research approach allows for interpreting meaning, but not correlating facts, which a positivist paradigm and quantitative research approach would allow; interpretivist and positivist research paradigms are different, but not contradictory, epistemological projects (Cohen, Manion, & Morrison, 2007; Karlsson, 2009; Monrouxe & Rees, 2009; Thanh & Thanh, 2015).

Qualitative research, "Seeks to understand people's experiences, the meanings they assign to those experiences, the psychosocial aspects of and language used in interpersonal interactions, and the factors that influence perspectives and interactions" (Ramani & Mann, 2016), or, simply put, "explain the elements of phenomena and their relationships" (Ringsted, Hodges, & Scherpbier, 2011). It aims to ask and answer the "how", "why", and "what is the nature of…" research questions, for example, how do lecturers think about and practice assessment, and why? The outcome of qualitative research in the field of education is to better understand behaviours, attitudes, environments and interactions, such as lecturer assessment practice (Ramani & Mann, 2016).

One might think that, in order to use Southern Theory as a conceptual framework, that a critical research paradigm should rather be employed, as Standpoint Theory/Epistemology traditionally

lies within the critical paradigm (for example, feminist research) (Fawcett & Hearn, 2004). However, while Southern Theory leads to the amplification of previously silenced voices, I argue that because these voices have been ignored or misrepresented, it is critically important to *understand* these individuals in their context, so, as understanding is the focus of this study, and not possible critical perspectives that arise, using an interpretivist research paradigm is appropriate.

Additionally, as Southern Theory is not the focus or core conceptual framework of this study, the focus is lecturer assessment practice (understanding behaviour) and not comparing North-South lecturer assessment practice, or specifically critiquing Northern lecturer assessment practice[7], so using an interpretivist paradigm over a critical paradigm remains a coherent choice (Bunniss & Kelly, 2010; Guba & Lincoln, 1982; Hunt, 1989; Maxwell, 2004a, 2004b; Monrouxe & Rees, 2009; Tavakol & Sandars, 2014a, 2014b). The use of Southern Theory as a conceptual framework in this study is as a lens within which to illuminate and magnify (highlight) an important aspect of lecturer assessment practice for comment (Bordage, 2009).

In light of HBT and the literature reviewed, evidence of conceptions as a critical factor in understanding and changing lecturer assessment practice emerged. While the main study seeks to understanding lecturer assessment behaviour, specifically what factors influence it, a related sub-study in this research endeavour is exploring lecturers' conceptions of assessment. To this end, exploring lecturers' conceptions of assessment, as a particular personal factor of importance, was the initial focus of this study's hybrid data collection, followed by identifying personal and contextual factors. In order to explore said conceptions, a Phenomenographic research methodology was used.

[7] This has also shaped the data collection process in terms of the focus of the interview questions: questions centred on behaviour (assessment practice), influencing factors and conceptions over critical comparison or power and social justice in assessment.

4.1.2. A Phenomenographic methodology was used in order to investigate the lecturers' conceptions of assessment

Within the interpretivist paradigm, two approaches are commonly used for the in-depth study of phenomena: phenomenology and phenomenography (Åkerlind, 2005; Stenfors-Hayes, Hult, & Dahlgren, 2013). Phenomenology, a more widely known and used methodology, seeks to describe the *common* meaning or ultimate understanding, the "*essence*" or *core* of a phenomenon, the "thing" that makes something what it *is* (Cohen et al., 2007; Creswell, 2007; Hopkins, Regehr, & Pratt, 2017; Larsson & Holmström, 2007). In contrast, Phenomenography is a methodology that seeks to the describe peoples' *varied conceptions* of a phenomenon or the qualitatively *different ways* in which people experience, interpret, understand and thereby make meaning of an aspect of reality or a phenomenon (Marton, 1981).

While both methodologies study human experiences and understandings of phenomena, considering a phenomenon from the participant's perspective, the focus or object of study differs: Phenomenology describes what is *common* to different forms of experience (what *is*) and Phenomenography describes a phenomenon through the *variation* in peoples' experiences (*understanding* what is) (Cibangu & Hepworth, 2016; Larsson, Holmström, & Rosenqvist, 2003; Limberg, 2000; Sjostrom & Dahlgren, 2002; Stenfors-Hayes et al., 2013; Tight, 2016). Phenomenology describes first order perspectives (describing the phenomenon: what is externally observed) and Phenomenography describes second order perspectives (how an individual experiences a phenomenon, their *conception* of what is, how something appears to a person in real life) (Holland, Middleton, & Uys, 2013; Marton, 1981; Moller, Fridlund, & Goransson, 2010; Munck, Sandgren, Fridlund, & Martensson, 2012b; Taylor, 2011; Woollacott, Booth, & Cameron, 2013). In this study, Phenomenology would ask, "What *is* assessment?" whereas Phenomenography asks, "What are *all the different ways* individuals think about and understand assessment?" (Munck et al., 2012b; Pihl, Fridlund, & Martensson, 2011; Stenfors-Hayes et al., 2012; Wahlstrom et al., 2001).

The aim of this sub-study is to identify and describe the diverse range of conceptions of assessment lecturers hold (Åkerlind, 2005; Boet, Sharma, Goldman, & Reeves, 2012; Harris, 2011; Marton, 1981; Stenfors-Hayes, Hult, & Dahlgren, 2013). These variations will assist in developing a comprehensive and explanatory model of lecturer assessment behaviour, as different ways of understanding may open up the possibility of new ways of doing, leading to the development of new practical theory, such as educational interventions (Larsson & Holmström, 2007).

The term "Phenomenography" was first used by Sonneman (1954), in the field of psychopathology, however, the Phenomenography methodology, as we understand it today, was developed in Sweden in the 1970-1980s, in the domain of learning by Marton and colleagues (Marton, 1981; Marton & Saljo, 1976). Traditionally, while Phenomenography was used in the domain of learning, today its use has spread globally to many research fields, including education and health sciences, (Ashworth & Lucas, 1998; Cibangu & Hepworth, 2016; Entwistle, 1997a; Limberg, 2000; Richardson, 1999; Stenfors-Hayes et al., 2013; Tight, 2016), and has been used in the global South, such as Australia (Åkerlind, 2003, 2017; Samuelowicz & Bain, 2002) and South Africa (Holland et al., 2013; Ojo & Booth, 2009; Taylor, 2011; Woollacott et al., 2013).

Phenomenographic analysis will be discussed in **Section 4.6.1.** (Åkerlind, 2005; Ashworth & Lucas, 1998; Larsson & Holmström, 2007; Marton, 1981; Richardson, 1999; Stenfors-Hayes et al., 2013). In short, Phenomenography posits that there are a limited number of ways in which individuals may experience and understand a phenomenon. These variations (conceptions) relate to one another in an organised hierarchy, "With each higher level encompassing those below it, and the highest level representing the most advanced or developed way of experiencing the phenomena" (Tight, 2016). These conceptions are grouped or unified into themes, called "categories of description", and their hierarchical relationships are displayed in an "Outcome Space" (Åkerlind, 2005; Ashworth & Lucas, 1998; Larsson & Holmström, 2007; Marton, 1981; Richardson, 1999; Stenfors-Hayes et al., 2013).

Critiques of Phenomenography include debates on initial research into ("surface" and "deep") approaches of learning (which are unrelated to this study and is not the concern of this thesis). More broadly, there are critiques associated with the interpretivist methodology in which Phenomenography falls (Webb, 1997a, 1997b). This methodological concern may be addressed through reflexivity and a felt responsibility by the research and participant as a co-constructors of knowledge (Ekeblad, 1997; Entwistle, 1997b). Issues of rigour will be detailed in **Sections 4.3.2., 4.3.3. & 4.4**.

More recently there has been debate in the field and use of Phenomenography, for example, distinguishing between "pure Phenomenography", "developmental Phenomenography" and "Variation Theory". Pure phenomenography has been described as, "How people conceive of various aspects of their reality, where the concepts under study are mostly phenomena confronted by subjects in everyday life rather than course material studied in school" (aligned with more traditional Phenomenography), whereas, developmental Phenomenography has been described as a research approach that is, "Shaped by the educational setting and therefore has intended outcomes, such as improving teaching and learning practices" (leading to action taking place) (Bowden & Walsh, 2000). Variation Theory, or the variation of learning theory, sometimes referred to as the "New Phenomenography", while rooted in Phenomenographic tradition, has progressed to focus on theories of awareness, theories of learning for educational value and the outcome of student learning (Åkerlind, 2017). Variation Theory differs from Phenomenography in that it focuses on identifying the pedagogically advantageous variant that benefits student learning (Marton & Pong, 2005; Pang & Ki, 2016; Pang & Marton, 2013; Rovio-Johansson & Ingerman, 2016).

However, others have argued that there is no real difference between Phenomenography and Variation Theory, as they shared the same underlying theoretical epistemological and ontological positionings (Åkerlind, 2017). In this study, I shall refer to and use "Phenomenography" as described in the original sense, because any comparison, critique or behaviour change is beyond the scope of this thesis.

4.2 Sampling strategies

4.2.1. Purposive and maximum variant sampling

In qualitative research, the number of participants required can vary greatly, averaging between five to sixty, depending on the research methodology employed (Creswell, 1998; Miles & Huberman, 1994; Morse, 1994, 1995). According to Phenomenographic literature, the average number of participants, as a guide for adequate sampling, ranged from ten and twenty participants, as this number was needed in order to develop a sufficient understanding and a deep and rich description of the phenomenon (Bowden & Walsh, 2000; Gillsjö, Schwartz-Barcott, Bergh, & Dahlgren, 2011; Holland et al., 2013; Larsson & Holmström, 2007; Moller et al., 2010; Munck, Sandgren, Fridlund, & Martensson, 2012a; Munck et al., 2012b; Ojo & Booth, 2009; Pihl et al., 2011; Stenfors-Hayes, Hult, & Dahlgren, 2011; Stenfors-Hayes et al., 2012; Wilhelmsson, Dahlgren, Hult, & Josephson, 2011; Woollacott et al., 2013). However, the number of participants is not of key importance, rather, the appropriateness and quality of the data collected (O'Reilly & Parker, 2012).

Any sampling strategy needs to address issues of data saturation and theoretical sufficiency (Morse, 1995, 2010, 2015a, 2015b; Morse, Barrett, Mayan, Olson, & Spiers, 2002). Data saturation refers to the idea of data fullness or depth and richness, meaning that enough data has been collected so that no new categories or themes emerge from further data collection and analysis (all aspects of the phenomenon have been obtained). The data is considered to be "saturated" or maximally filled in a category or theme (Bowen, 2008; Dicicco-Bloom & Crabtree, 2006; Guest, Bunce, & Johnson, 2016; Morse, 1995, 2015b; Morse et al., 2002; O'Reilly & Parker, 2012). Morse (2015b) describes the full scope and process of data saturation as, "The phenomenon becomes stronger, more evident, more consistent, more cohesive, and more mature, research becomes saturated, and the researcher becomes certain." Data saturation means that a phenomenon and all its aspects are deeply and fully explored, and common essential characteristics or themes have been consistently found or replicated across data from

several different participants (they are prevalent and stable), and all major variations are identified and then incorporated into the developing theory (Bowen, 2008; Guest et al., 2016; Morse, 2015b). This is achieved through more than just appropriate sampling, but, also through multiple independent researchers analysing the data, a spiralling back and forth between the researcher and data from participants, and the keeping of a cumulative and detailed analysis audit trail (providing evidence for each analytical step in the process towards achieving data saturation), eventually leading to the development of a more confident and certain theory (Bowen, 2008; Guest et al., 2016; Morse, 1991, 2015b).

Theoretical sufficiency refers to, "the stage at which categories seem to cope adequately with new data without requiring continued extensions and modifications," meaning that the developing theory has a sufficient foundation of evidence to support its propositions, including, rich and deep descriptions, replicated data, prevalent and stable themes (Cohen et al., 2007; Morse et al., 2002; O'Reilly & Parker, 2012; Petty, Thomson, & Stew, 2012; Schut et al., 2018; Varpio et al., 2017). Varpio et al. (2017) makes the important point that theoretical sufficiency is about having "enough" evidence for the claim that is being made, it does not have to be exhaustive. Adequate sampling, data saturation and theoretical sufficiency all contribute to rigour in qualitative research (discussed in **Section 4.4**.). For adequate, appropriate and purposeful sampling of participants, purposive and maximum variant sampling were used in this study.

Purposive sampling refers to the selection of participants that have experienced the event of interest or are most qualified to provide rich perspectives, in other words, these individuals meet the predetermined, relevant criteria (Ramani & Mann, 2016). In this study, lecturers from final year/s in medical programmes who have experience practicing assessment in Southern contexts.

Maximum variant sampling, a type of purposive sampling, is the selection of respondents from different contexts to increase the likelihood that the findings will reflect differences or varied perspectives (Bowden & Walsh, 2000; Creswell, 2009; Tavakol & Sandars, 2014b). This is in line

with Phenomenography in seeking to describe the range of conceptions (I am arguing that by sampling from diverse participants in diverse contexts, it is more likely that a varied range of experiences and understandings, thus conceptions, are found) (Åkerlind, 2005; Holland et al., 2013). It much also be noted that the goal of Phenomenography is not to uncover the extent (intensity, distribution, coverage or frequency) to which a phenomenon exists but rather the range of variation (Stenfors-Hayes et al., 2011, 2012). In this study, diversity of participants was defined by personal (demographic) characteristics, such as age, gender, level of qualifications or training received, years of experience and discipline, and contextual characteristics, including nationality, HEI and country of practice. Personal and contextual diversity were selected as they align with HBT and the influencing factors identified in the literature (**Section 3.3.**).

Having adopted Southern theory as part of the conceptual framework for this study, sampling was limited to three Southern settings: two from South Africa and one from Mexico. While "Southern", in Southern Theory, is conceptual and not geographical, in terms of "contextually diverse" sampling, these countries may also be described as "developing" or "resource-constrained"[8] (see **Table 4.1.:** of interest to this study is not only economic wealth but also the crude measures of health care system efficiency, such as maternal mortality and infant mortality. For the purposes of the study, these issues are important as they help frame the context in which HPE takes place. These are rough indicators of the burden of disease the healthcare system – within which clinical HPE takes place and in which clinical lecturers must operate must cope with). Additional descriptors of "contextual diversity", for example, colonial histories, cultures and major religions are outlined in **Table 4.2.** (see **Appendix 4.1.**).

[8] Generally, "developing" or "resource-constrained" countries, sometimes referred to as low- and middle-income countries, have lower per capita GDP and higher maternal and infant mortalities than "developed" and "resourced-rich" contexts in the global North or West (The World Bank).

Table 4.1.: Sampling sites statistics: Developed versus developing countries statistics in which data collection took place (WHO, 2018).

Country	**Per capita Gross Domestic Product (GDP)** (US$)	**Doctor: Patient** (physician per 1000 population)	**Maternal mortality** (deaths per 100 000 live births)	**Infant mortality** (deaths per 1000 live births)	**Life expectancy** (total years)
Developed (resource-rich) countries:					
United States of America	$ 59 160	2.6	14	6.5	78.5
Netherlands	$ 46 910	3.5	7	3.8	81.6
Australia	$ 51 360	3.5	6	3.7	82.5
United Kingdom	$ 40 600	2.8	9	4.3	81.2
Developing (resource-constrained) countries:					
Country	**Per capita Gross Domestic Product (GDP)** (US$)	**Doctor: Patient** (physician per 1000 population)	**Maternal mortality** (deaths per 100 000 live births)	**Infant mortality** (deaths per 1000 live births)	**Life expectancy** (total years)
Mexico	$ 16 110	2.2	38	14.6	77.3
South Africa	$ 5 430	0.8	138	43.3	63.4

Sampling was also informed by the concept of transferability. By sampling from multiple, varying contexts, there is a greater potential that other researchers and practitioners may find the results and theory generated in this study to be transferable than would be the case if sampling were limited to a single setting (related to rigour in **Section 4.4.**) (Bowen, 2008; Korstjens & Moser, 2018; Malterud, 2001; Miles & Huberman, 1994; Petty et al., 2012).

4.2.2. Pragmatic delimitations

Sampling was pragmatically delimited to lecturers who assess in the final year/s of medical programmes, from diverse disciplines and Southern settings were included in the study.

Lecturers in medical programmes were selected because there is a paucity of information and theory around HPE lecturer assessment practice, especially in the global South. Moreover, lecturers in medical programmes are generally clinicians who possess little or no educational or assessment training, indicating a potential need for faculty-training.

Lecturers from the final year/s of medical programmes were selected as assessments at this stage of the medical programme are of a high stakes nature because serious consequences exist in terms of certifying or licensing a student as competent and safe for public practice.

When speaking of lecturers, I am specifically referring to sampling from course conveners, or course coordinators, when possible, because in the global South assessment is generally the responsibility of a single individual as opposed to an assessment committee or team. The course convenor is therefore a key role player in assessment for they are likely to be in charge and oversee assessment practice in their clinical rotation. This makes the individual lecturer an important unit of intervention in any attempts to change assessment practice.

4.3. Methods

In this study, data collection took place in three stages: a pilot stage, the first round of validation (the first part of the confirmatory study) and a second round of validation (the second part of the confirmatory study leading to the final results and conclusion) (see **Figure 4.1.**). Semi-structured interviews took place during each stage with one qualitative semi-structured research interview per individual, with the request for potential follow-up interviews if necessary. In total **31 participants** were interviewed: 12 lecturers were interviewed from a single South African HEI for

the pilot study (labelled "SA1-12"), followed by 19 interviews during the confirmatory study stage, which consisted of 6 additional interviews at a different South African HEI ("SA13-18), and 13 from a Mexican HEI ("MX1-13"). All interviews were audio-recorded and transcribed verbatim, resulting in 231 000 transcribed words. Field notes were also taken either during or immediately after the interviews, noting, for example, observed and heard nonverbal cues and communications (tone, gestures, facial expression, reactions) which would not be inferable from transcribed text. Reflexive notes detailing personal feelings, thoughts, perceived rapport were added to these field notes, as well as any initial ideas, themes, similarities or differences that seemed to be potentially important (Denscombe, 2010; Knox & Burkard, 2009; Whiting, 2008).

The interviews ceased when data saturation and theoretical sufficiency were reached, which occurred around the third stage of data collection. As data analysis was iteratively performed, I was aware of the developing Outcome Spaces and themes, before each subsequent interview. During the third stage of data collection and analysis in Mexico, the sixth interview ("MX6"), data saturation was reached as no new constructs (conception categories nor elements, or personal and contextual factors) emerged. Indeed, the following six interviews confirmed data saturation. Theoretical sufficiency was felt in the quality and coherence of the data collected, and will be demonstrated in the quotations provided.

The interview questions (see **Appendix 4.2.**) used in this study were previously used by the researcher's supervisor in a related study (demonstrating the validity of the questions). After each cycle of data collection, as new conceptions emerged, the interview questions were adapted as needed. The interview questions acted only as a guide and not a set agenda. For example, "identity", which was not present in the first phase interview questions, emerged during the first interview of the second phase of the study. During subsequent interviews in phases two and three, participants were asked what they identified themselves as, whether or not they saw themselves as clinicians or felt a stronger loyalty towards calling themselves an educator.

The researcher of this study piloted, not for data collection purposes, these interview questions and the semi-structured interview process with lecturers in an unrelated field for experience and training before beginning formal data collection.

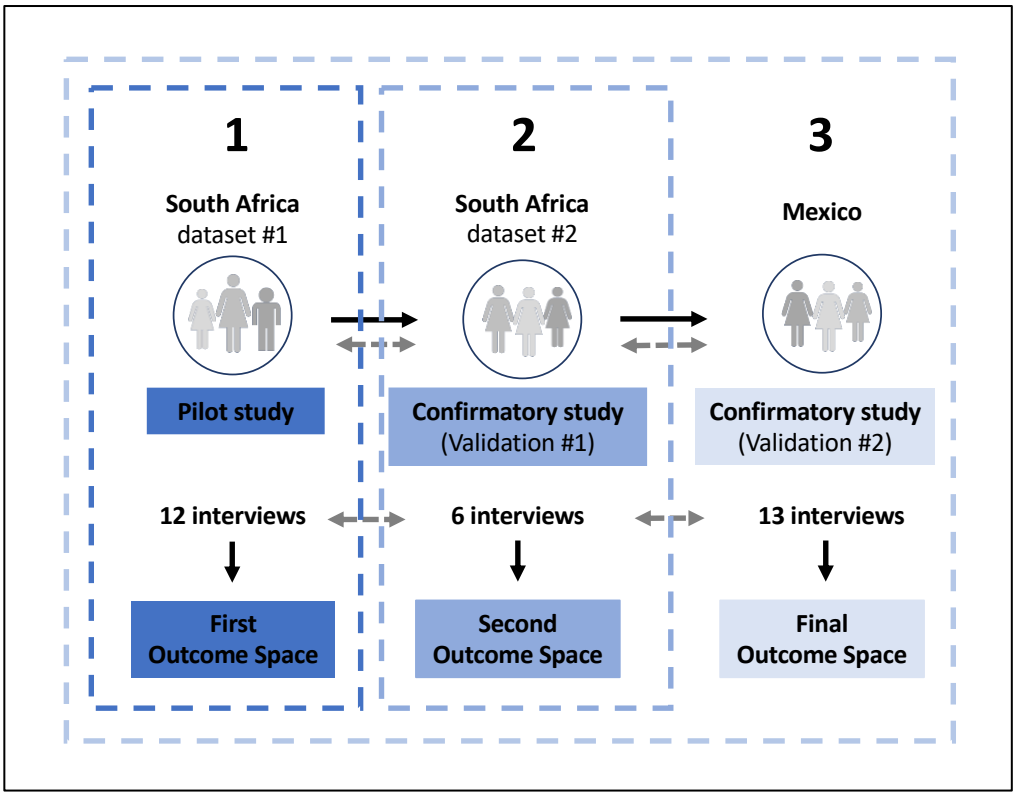


Figure 4.1.: Data collection: an overview. Data collection occurred in three stages: (1) a pilot study was conducted at an South African HEI, resulting in 12 interviews, and data was then analysed inductively before (2) interviews at a second, different South African HEI took place, resulting in an additional 6 interviews. This was followed by inductive data analysis and comparison to the first dataset for similarities, differences and validation of the proposed conclusions and developed models. These stages were followed by (3) further interviews at an additional Southern site, Mexico, resulting in 13 interviews, with deductive analysis taking place for further model validation.

4.3.1. Data Collection: Qualitative Research Interviews

The purpose of an interview is to collect data or gather information, build knowledge and understand experience, through interviewer and interviewee interaction (Cohen et al., 2007; Creswell, 2007, 2009; Dicicco-Bloom & Crabtree, 2006; Knox & Burkard, 2009; Mann, MacLeod,

Cleland, & Durning, 2015; McGrath, Palmgren, & Liljedahl, 2018). Interviews also serve the function of exploring more complex and subtle phenomena, such as seeking to gain insight and understanding in respondents' beliefs, emotions and experiences, as opposed to simply collecting factual information, for which the instrument of questionnaires would serve as a more cost-effective option. In contrast, during the interview process, respondents are *participants* in meaning making, rather than just conduits from which information is retrieved (Denscombe, 2010; Dicicco-Bloom & Crabtree, 2006; Knox & Burkard, 2009). The challenge with interviews is how to simultaneously make respondents feel safe and comfortable while deeply exploring experiences to yield rich and meaningful data with a relative stranger (Knox & Burkard, 2009; McGrath et al., 2018). How this was achieved is detailed in the section on reflexivity (see **Section 4.3.3.**). For this study, qualitative research interviews were chosen as the method of data collection for they were the most appropriate way of answering the questions I was exploring and they aligned with the research methodologies employed.

Qualitative research interviews may be described in many ways: a dialogue, a conversation with a purpose, and the co-construction of knowledge resulting from the interpersonal relationship between the interviewer and participant. Kvale (1992) describes interviews as, "A joint endeavour where egalitarian partners, through conversation, search for true understanding and knowledge." However, depending on the researchers' epistemological and ontological perspectives, qualitative research interviews may be understood in different ways. For example, the neo-positivist views interviews as a tool to gather facts ("A context-free truth about objective reality producing relevant responses, with minimal bias"), the romanticist sees interviews as human encounters ("(That) reveal their authentic experiences by establishing rapport, trust and commitment between the interviewer and interviewee"), whereas, the localist perspective claims that objective interviews do not exist as information gathered cannot be separated from the context in which it was gathered (Qu & Dumay, 2011). For the purposes of this study, interviews were viewed from the localist perspective, as social constructions of situated accounts, because the research questions of this investigation focused on exploring meaning from diverse perspectives (persons and contexts) (Qu & Dumay, 2011).

Potential critique of qualitative research interviews, generally by researchers operating in positivist research paradigms, includes interviews being seen as unscientific, subjective, untrustworthy, unreliable and invalid. Kvale (1992) responds to these criticisms by asking the definition of those terms: what does it mean to be scientific? What is objectivity? What is trustworthiness? What is validity? Reflexivity (see **Section 4.3.3.**) speaks to issues of "subjectivity" and bias. When considering the interviewer as the research tool, it may be said that while different interviewers may arrive at different decisions and conclusions during the research process, the principles of hermeneutics show that multiple interpretations are legitimate and can actually be a strength and not a weakness in terms of data richness (Kvale, 1992). Moreover, an inherent advantage of interviews is that they possess communicative and pragmatic validity, for, as information is being gathered it may be directly checked for clarity, accuracy and relevance (Denscombe, 2010; Kvale, 1992). The interview transcripts may also be checked by the respondents ("member checking") to ensure further rigour (see **Section 4.3.3.**).

Another point of contention is the use of leading questions, however, the important issue is not the leading questions themselves but, rather, to where do they lead – new knowledge? As the aim of qualitative interviews is to explore, reflect the nature of a phenomenon and develop theory, then trying to force positivist parameters, which are better suited to hypothesis testing, is inappropriate (Kvale, 2006). Due to the small sample sizes of qualitative studies the results are not generalisable, yet, again, as the goal of qualitative research is not to generalize findings but, rather, understand a phenomenon richly and deeply, locally contextualizing knowledge, then, through this, findings may be transferable from one setting to another (Kvale, 1992).

There are different types or formats of interviews: structured, semi-structured, unstructured, one-to-one, group and focus groups interviews, which may take place either in person or virtually, for example, online (Denscombe, 2010). The type of interview selected depends on the purpose the researcher wishes to accomplish. For instance, a structured interview is often used to collect quantitative data, due to the highly-controlled and standardised nature of the interview

in terms of predetermined questions and limited, pre-coded answers given for respondents to choose from (Denscombe, 2010; Dicicco-Bloom & Crabtree, 2006). An unstructured interview allows the respondent to use their own words, develop ideas and pursue their own thoughts, and is often used in researching complex, social issues (Denscombe, 2010; Dicicco-Bloom & Crabtree, 2006). For this study, the format of semi-structured interviews was used.

4.3.2. Semi-structured interviews

Semi-structured interviews steer information in a particular direction, or theme, but also allows for open answers and discussion around said direction (Denscombe, 2010; Dicicco-Bloom & Crabtree, 2006; Knox & Burkard, 2009; Qu & Dumay, 2011; Ramani & Mann, 2016; Whiting, 2008). This takes place through the researcher asking questions, from a list of pre-selected questions around a specific topic, yet, it also provides an opportunity for the participant to give an answer in their own words, as opposed to choosing a limited and pre-selected answer (Cohen et al., 2007; Tavakol & Sandars, 2014b). Cohen et al. (2007) state, "The semi-structured questionnaire (interview) sets the agenda but does not presuppose the nature of the response." The open-ended nature of the questions are used to invite honest and personal comments, which allows for unknown answers and explanations of a complex issue, and for the participant to elaborate on points of interest. A further advantage of open-ended questions is their flexibility, allowing the interviewer to clear up misunderstandings or probe to a greater depth, establish rapport, encourage cooperation and ultimately provide a truer picture of what the respondent really means (Cohen et al., 2007; Denscombe, 2010; Qu & Dumay, 2011; Whiting, 2008). Indeed, the while there was a list of questions I desired to work through with the participants interviewed (and ensured that all areas of interviewed were covered in each interview), the interviews themselves rarely followed a prescribed order, because, as participants raised new or interesting points, the interview naturally followed and explored these new directions, reflecting the iterative nature of qualitative data collection and analysis. Semi-structured interviews are also the most commonly used method of data collection in Phenomenography, because they encouraged deeper reflection and a full exploration of a participant's understanding of the

phenomenon, which reaffirms their use in this study (Entwistle, 1997a; Gillsjö et al., 2011; Qu & Dumay, 2011; Stenfors-Hayes et al., 2011; Tight, 2016; Wahlstrom et al., 2001; Whiting, 2008).

In line with the aims and objective of this study, interview questions focused on deeply exploring lecturer assessment practice in diverse Southern settings, specifically seeking to identify factors that shape their assessment practice, and exploring lecturers' conceptions of assessment as a sub-study (interview questions can be found in **Appendix 4.2.**). Due to the iterative nature of qualitative research, questions in subsequent interviews changed in light of the data previously collected and analysed: confusing and ineffective questions were removed, other questions were altered to better uncover necessary information, and new questions were added as new constructs emerged (Dicicco-Bloom & Crabtree, 2006; Knox & Burkard, 2009; McGrath et al., 2018).

For example, in initial interviews, questions asked were, "What is the purpose of assessment?" and "Why do you practice it?" It became apparent that a potentially better approach to ease participants into the interview process was to ask a more familiar and comfortable question such as, "Please describe your assessment practice to me – what do you do?" Participants' concrete descriptions of their assessment practice then acted as a foundation on which more abstract, conceptual questions could be ask, such as, "Why do you do that?" or "Can you explain your reasons for doing that?" or "How did you come to that decision?" As data was collected, and preliminary results began to emerge, probing questions also became more specific, for instance, "A colleague mentioned a compulsory HPE training course at your HEI, what are your thoughts on it?" and "Did this course influence your assessment practice? How so?" or "Did you find the training course valuable, why or why not?" or "In another interview it was raised that clinical workload responsibilities limit potential time spent on assessment tasks, is this true for you? How? Why or why not?" While the interviews began broadly, allowing for open responses, towards the end of the interviews, particular factors were explored, to be confirmed or disputed, which also assisted in identifying when data saturation and theoretical sufficiency were reached.

Semi-structured interviews, in this study were, conducted in a one-to-one format. The majority of the interviews (30/33 = 91%) took place in person (face-to-face), with the remaining (3/33 = 9%) interviews done virtually (online, via audio-visual Skype calling). A one-to-one format was selected in order to establish rapport and deeply explore an individual's experiences and understandings, which would be more difficult to do in a larger group setting (Denscombe, 2010; Dicicco-Bloom & Crabtree, 2006).

While the majority of interviews were conducted in person, there are benefits to conducting interviews online. Advantages include, decreased cost and increased scale (overcoming geographical and financial barriers, increasing access to participants, both locally and internationally), leading to an increase in diversity and representation (Chesterman, Lamanna, Kalamatianou, & Rosenstock, 2017; Couper, 2005; James & Busher, 2016; Janghorban, Latifnejad Roudsari, & Taghipour, 2014; O'Conner, Madge, Shaw, & Wellens, 2008; Weller, 2017). While online interviewing allows for greater reach and access to more participants, it could be argued that the requirement for internet access is a bias towards developed countries, thereby further marginalizing those whose voices are not heard and discriminating against varied levels of technical competence (O'Conner et al., 2008). However, recent statistics show mobile phone penetration increasing rapidly in developing countries: in Africa, 43% of the population were connected via mobile phone in 2010 with a predicted increase to 54% by 2020, indicating that it is possible for even the poor in developing countries to increasingly own mobile phones through low-cost phones and prepaid subscription options and the potential of having internet access, opening up the possibility of conducting online interviews via smartphones if necessary (Aker & Mbiti, 2010; Chesterman et al., 2017; GSMA, 2016; ITU, 2016; Kreutzer, 2009).

A further advantage of online interviews is its increased flexibility and convenience, which could translate into more participation. Participants have also reported feeling less pressure in online interviews than in face-to-face interviews, because of the felt "distance", allowing participants to feel "safe" by conducting the interview in a familiar and more private space, such as at home

(Chesterman et al., 2017; Couper, 2005; James & Busher, 2016; Janghorban et al., 2014; Jenner & Myers, 2018; O'Conner et al., 2008; Weller, 2017).

There is debate as to whether or not face-to-face and online interviews leads to a difference in results, especially in establishing rapport between the researcher and the participant (Lo Iacono, Symonds, & Brown, 2016). Numerous studies have shown no significant difference, rather, a comparable level of rapport was felt in both face-to-face and online interviews (Janghorban et al., 2014; Jenner & Myers, 2018; Weller, 2017). As aforementioned, some participants preferred online interviews because of the "distance" which did not lead to a loss of intimacy, but, rather, removed the "pressure of presence", allowing these participants to more freely and deeply communicate (Weller, 2017). While one study reported a decrease in data collected (total word length of the interview), the format of the online interview was synchronous text-based (instant messaging) interviewing and not verbal interviewing (Jowett, Peel, & Shaw, 2011). Furthermore, one study showed no decrease or difference in data collected (total time of the interviews) between face-to-face and online audio-visual interviews (Jenner & Myers, 2018).

4.3.3. The importance and practice of reflexivity

An important consideration when using interviews as a data collection method is the interviewer effect. Participants respond differently (the willingness, extent and degree of honesty to which information is divulged) depending on how they perceive the interviewer and, in particular, their age, gender and ethnic origins or personal identity, as well as social status, cultural background, language, economic status, educational and professional levels, which may lead to unequal power dynamics (Denscombe, 2010; Knox & Burkard, 2009; Verdonk & Abma, 2013). While personal attributes may not be changed, their affect may be limited throughout creating a safe space within which the interview may be conducted, having the interview in an open, safe, comfortable and honest environment, being punctual, polite, sensitive, receptive, respectful and neutral to establish a rapport and encourage maximal participation (Denscombe, 2010; Dicicco-Bloom & Crabtree, 2006; McGrath et al., 2018).

One way in which an interviewer develop a degree of neutrality is through reflexivity. Reflexivity has been described in various ways: “Outing the researcher” (Finlay, 2002b), “Ongoing self-critique and self-appraisal” (Koch & Harrington, 1998), “Ways of being” (as opposed to ways of doing) (Attia & Edge, 2017), “Managing identity performance” (Bishop & Shepherd, 2011), and, “(An) internal dialogue and constant scrutiny” (Jootun, McGhee, & Marland, 2009). Reflexivity may then be defined as an ongoing process of thoughtful self-analysis and reflection, to make the interviewer increasingly aware of, and seek to, minimise potential influencing factors, such as age, gender, ethnicity, personal prejudice and bias (Attia & Edge, 2017; Darawsheh, 2014; Finlay, 2002b; Korstjens & Moser, 2018; Probst, 2015; Verdonk & Abma, 2013). Koch and Harrington (1998) put it well, “How do we study others without studying ourselves?”

It is important for researchers to be reflexive, particularly in qualitative research, where the researcher is the qualitative tool through which a study is conceived, literature reviewed, data collected and analysed, because, the potential impact of the researcher on these processes must be acknowledged, interrogated and disclosed in a detailed and descriptive manner (Bishop & Shepherd, 2011; Engward & Davis, 2015). Reflexivity is part of the entire research process from beginning to end, “How does who I am, who I have been, who I think I am, and how I feel affect data collection and analysis – that is, an acceptance and acknowledgment that “*how* knowledge is acquired, organized, and interpreted is relevant to *what* the claims are”” (Pillow, 2003). The researcher, and audience, need to be aware of possible influencing factors, personal beliefs, biases and experiences, social, cultural, political and contextual factors (Bishop & Shepherd, 2011; Engward & Davis, 2015). Indeed, “We emphasise the importance of the researcher consciously stepping back from action in order to theorise what is taking place, and also stepping up to be an active part of that contextualised action” (Attia & Edge, 2017). Reflexivity is a component of ethical research, moving towards honesty, transparency, integrity, trustworthiness and rigour (Bishop & Shepherd, 2011; Darawsheh, 2014; Guillemin & Gillam, 2016; Koch & Harrington, 1998). However, two things must be noted: firstly, the goal of reflexivity in disclosing subjectivity is not objectivity and detachment (as required in a positivist paradigm),

for subjectivity, when used with critical awareness, can be an advantage in qualitative research (Jootun et al., 2009; Koch & Harrington, 1998). Secondly, true reflexivity can never be totally achieved for human beings are, by nature, subjective, self-conscious, socially-constructed and continuously evolving (Finlay, 2002b).

Reflexivity may be practiced through various ways, such as critical awareness and reflection in general, for instance, the keeping of a research journal (Cohen et al., 2007; Creswell, 2007, 2009; Darawsheh, 2014; Koch & Harrington, 1998; Korstjens & Moser, 2018; Mann et al., 2015; McMillan, Cleland, & Durning, 2015; Probst, 2015; Tavakol & Sandars, 2014b; Whiting, 2008), or, more specifically, through explicit narrative reconstruction (the creation of "self-narratives": personal stories) (Bishop & Shepherd, 2011), or by observation and interaction with colleagues, collaborating and corroborating the trustworthiness of the data and interpretations (Attia & Edge, 2017; Finlay, 2002a; Konradt, Otte, Schippers, & Steenfatt, 2016; Probst, 2015). Being reflexive is not about ticking the reflection "box", rather, as Attia and Edge (2017) say, it is a way of being over a way of doing, "Reflexivity does not prescribe specific types of responses to research situations, rather, it is a sensitizing notion that can enable ethical practice to occur in the complexity and richness of social research" (Guillemin & Gillam, 2016).

In light of the conceptual positioning of this study (Southern Theory) and maximum variant sampling strategy (of demographically and contextually diverse participants from different contexts), it was also important for the researcher to declare their cultural and political context, and its related power dynamics (Engward & Davis, 2015; Finlay, 2002a, 2002b; Koch & Harrington, 1998). This lead to a critical awareness of the non-neutral and dominant education theories and methodologies being used in this study, and how I needed to guard against implicitly reproducing "Western" or "Northern" and "Whiteness" based norms (Mruck & Breuer, 2003; Verdonk & Abma, 2013). Reflexivity, like Southern Theory, is concerned with the positioning (and power) of the researcher and participant, and how knowledge is socially produced (and not reproduced or exploited): is research being conducted *on* or *with* participants? (Pillow, 2003).

Reflexivity was practiced through interview preparation in gathering background information on the sampling sites and speaking to local collaborators before beginning interviews (asking for their insight or recommendations, to prepare and familiarise myself with the context and people) (see **Tables 4.1.** and **Appendix 4.1.**), developing interview skills through the practice of pilot interviews, writing assumptions and critical reflections in a research journal, being aware of potential power inequalities (age, gender, ethnicity, education, profession, nationality, culture, religion), practicing sensitivity, open-mindedness and respect during the interviews, keeping field notes immediately after interviews, reading through and discussing transcripts with other independent researchers, interrogating interpretations and general rigour.

Indeed, I kept a research journal throughout the entire research project, have declared "self" (see **Appendix 4.3.**), critically discussed findings with independent researchers and practiced rigour throughout (see **Section 4.4.** below). In my research journal, subjectivity was acknowledged, examined and declared, and, in line with rigourous qualitative research practice, detailed thinking, explanations, rationale and evidence given for decisions made and conclusions drawn were written there (Engward & Davis, 2015).

I found explicitly stating my understanding, feelings and conclusions drawn after each interview, helpful in identifying and mitigating possible clouding emotions and interfering biases. For instance, dealing with issues of age and gender in interviewing generally older, male clinicians, making notes on the perceived rapport and detailing possible interactions and experiences that may sway my impressions of the individual and what they had said, were detailed in my research journal and interrogated. The research journal was also useful in tracking the development and maturation of my thinking and understanding during the analytical processes. Periodically I would read back through my research journal to remind myself of how ideas evolved and deepened over time.

I also found it helpful in beginning the interviews with a brief statement of introduction to put the participant at ease. As I received some surprised expressions and exclamations of, "You're so

young!" upon introduction, I would start by sharing my educational background and qualifications. This provided participants with a sense of openness and safety (a non-threatening space). I would tell participants that I was not a clinician, nor an educationalist, merely a basic scientist who made the jump to educational research a couple years ago. This dampened some of the nerves and pressure felt for both myself and the participant. I believe that a rapport was established in each interview and the discussions were honest and productive. At first, some participants may not have taken the interview seriously, but, as it progressed, and thought-provoking questions were asked, and their thinking interrogated, I could see the change in their attitude towards myself and the research project, genuinely thinking deeply and freely sharing their views and opinions. As data collection progressed, and my confidence grew (in terms of my knowledge, skills and sense of certainty towards the developing results), I began to thoroughly enjoy myself, despite the complex process of interviewing. When I reflect on my role in the processes of data collection and analysis, I do believe that a sufficient degree of neutrality and reflexivity was achieved, allowing the findings of this study to speak for themselves.

4.3.4. Establishing rigour

Critiques and disadvantages of interpretivist and qualitative research methods (as opposed to positivist and quantitative research methods) have been described by many, such as clinicians questioning the legitimacy of qualitative research, perceiving it to be invalid and unreliable: the "subjective" nature of findings that are open to researcher bias, and its limits to generalisability (Albert, Laberge, Hodges, Regehr, & Lingard, 2008).

These concerns, and others, have been addressed in **Section 4.3.2. & 4.3.3.** and will be further detailed below, for, qualitative research approaches come from a long tradition of sociological research tradition and is accepted as an established research approach in HPE (Britten, 2005; Bunniss & Kelly, 2010; Guba & Lincoln, 1982). This research project was also based on, and supported by, theoretical perspectives and consulted literature (Albert, 2004; Bordage, 2009; Cook et al., 2008; Prideaux & Bligh, 2002). Principles of rigour applied in this study address many of the concerns (Dornan, Peile, & Spencer, 2009; Dornan & Spencer, 2008; Eva, 2009).

In qualitative research a number of standards and quality assurance criteria (see **Table 4.2.** which provides details on how each component may be achieved) need to be taken into account in order to ensure meaningful results and defensible conclusions, including trustworthiness, credibility, confirmability, dependability, utility (applicability), transferability and ethically-guided practice (Anderson, 2010; Bowen, 2008; Guba & Lincoln, 1981; Koch & Harrington, 1998; Korstjens & Moser, 2018; Levitt, Motulsky, Wertz, Morrow, & Ponterotto, 2017; Malterud, 2001; Morse, 2015a; Morse et al., 2002; Moss et al., 2009; O'Brien, Harris, Beckman, Reed, & Cook, 2014; Petty et al., 2012; Santiago-Delefosse, Gavin, Bruchez, Roux, & Stephen, 2016; Varpio et al., 2017).

In this study, rigour was established through the appropriateness and alignment of the study, from its research questions and conceptual frameworks to its data collection and sampling strategies, described in detail, with data collection taking place across multiple diverse sites, until data saturation and theoretical sufficiency were reached, and through the practice of reflexivity (Ashworth & Lucas, 1998; Holland et al., 2013; Levitt et al., 2017; Malterud, 2001; Moller et al., 2010; O'Brien et al., 2014; Pihl et al., 2011; Santiago-Delefosse et al., 2016). Data collection consisted of using previously validated questions, practice interviews for training, and interviews were conducted, transcribed and analysed by the same researcher (Moller et al., 2010; Pihl et al., 2011). Data saturation and theoretical sufficiency were reached through iterative data collection and analysis, careful and repeated readings transcript readings, and robust discussions with additional, independent researchers (Åkerlind, 2005; Holland et al., 2013; Moller et al., 2010; Pihl et al., 2011; Ramani & Mann, 2016; Schut et al., 2018). Results given are supported by thick and rich descriptions and numerous, detailed quotations (see **Chapter's 5, 6** and **7**) (Munck et al., 2012b; Petty et al., 2012; Santiago-Delefosse et al., 2016; Sjostrom & Dahlgren, 2002). Constant comparison took place and negative cases or competing explanations were considered, contributing to credibility, trustworthiness, confirmability and dependability (Anderson, 2010; Morse, 2015a; Schut et al., 2018). Support for the findings of this study was garnered across

different settings (Åkerlind, 2005; Ramani & Mann, 2016; Ringsted et al., 2011; Santiago-Delefosse et al., 2016; Watkins et al., 2005).

Table 4.2.: Rigour in qualitative research: an overview.

Criteria	How is it achieved?	Reference
Trustworthiness: credibility, confirmability and dependability	• Prolonged engagement or persistent observation • Participant involvement or member checking or peer debriefing • Coherence with paradigm • Triangulation • Audit trail or clear detailing of research process • Constant comparison • Theoretical sufficiency • Negative cases • Reflexivity	• Anderson (2010) • Bowen (2008) • Guba and Lincoln (1981) • Korstjens and Moser (2018) • Malterud (2001) • Miles and Huberman (1994) • Morse (2015a) • O'Brien et al. (2014) • Petty et al. (2012) • Santiago-Delefosse et al. (2016) • Varpio et al. (2017)
Utility (applicability) and transferability	• Transparency in research process: explicit details • Purposive and sufficient sampling • Contextualisation • Thick and rich descriptions • Relevant and meaningful findings • Coherence	• Bowen (2008) • Guba and Lincoln (1981) • Korstjens and Moser (2018) • Levitt et al. (2017) • Malterud (2001) • Miles and Huberman (1994) • Morse (2015a) • O'Reilly and Parker (2012) • Petty et al. (2012) • Santiago-Delefosse et al. (2016)

Interestingly, Moss et al. (2009), proposes that for educational research to be considered of high quality and rigour, it needs to be educationally imaginative and socially significant, possible challenging or disrupting current standards or practices. In this study, there is a paucity of

information regarding lecturer assessment practice, especially in the global South, leading to both educationally significant findings and the contribution to powerful, socially just Southern Theory.

While reflexivity has already been described as a component of rigour, and as an example of ethical practice, ethical considerations in general, will be discussed below.

4.3.5. Ethical considerations

In order to address the ethical principles of autonomy (informed consent, protecting anonymity and confidentiality, right to withdraw), nonmaleficence and beneficence (favourable risk:benefit ratio) (Cohen et al., 2007), several steps were taken, detailed below.

Ethical clearance was obtained from the Human Research Ethics Committee (HREC) at the Faculty of Health Sciences of UCT (HREC reference: 689/2017), and from participating institutions, if needed. It was not anticipated that the study would raise ethical issues as the educational value was high and the risk:benefit ratio favourable (Cohen et al., 2007).

Potential participants were clearly informed about the nature of the study, the purpose of the interview, as well as relevant details (such as the format and length of the interview), and then invited to participate, whereby voluntary, informed and written consent was elicited (see **Appendix 4.4.** for consent form) (Cohen et al., 2007). Participants were informed as to the purpose of the recorder (for later transcribing of interviews) and permission was obtained before proceeding with the interview (Whiting, 2008). Due to the nature of this study, in deeply exploring lecturer conceptions of assessment and how various factors influences assessment practice, any possible, unintended harm or exploitation were minimized, especially as participants discussed their experiences (Dicicco-Bloom & Crabtree, 2006; Ramani & Mann, 2016). Potential participants were informed that they would face no penalty for not participating, and if they did participate, they maintained the right to withdraw from the study at any time

without prejudice or penalty (Cohen et al., 2007; Whiting, 2008). Participants were also allowed to ask any questions for clarification throughout the interview. Participants were treated sensitively and with respect at all times, individual privacy was upheld and responses will remain confidential without intrusive threats, stigmatisation, incrimination or revelations disclosed of private, stressful or sacred areas of the respondents in line with the ethical principle of non-maleficence (Callahan, Hojat, & Gonnella, 2007; Cohen et al., 2007; Dicicco-Bloom & Crabtree, 2006; Ramani & Mann, 2016; Whiting, 2008).

Data from this study, anonymised, has been securely stored on a password protected computer accessible to a single researcher (the author of this thesis). Participants will be informed as to the final results of this study in the form of presentations and publications, allowing respondents to see their contribution to this study (Cohen et al., 2007; Dicicco-Bloom & Crabtree, 2006). All data collected (recordings, transcripts, notes) are, and will continue to be, securely stored, accessible only by the researchers of this study. Data will remain in secure storage for five years and then destroyed.

4.3.6. Data analysis

As the outcomes of qualitative research in an interpretivist paradigm are developing new, or expanding, concepts, clarifying and understanding complexity, providing insights that refine knowledge and change behaviour, and developing theory, which is in line with the research questions of this study, the qualitative research process it iterative (Major & Savin-Baden, 2011; Peshkin, 1993) This means that data collection and analysis occurred simultaneously, and, that preliminary data affected subsequent data collection and analysis, which led to a spiralling back-and-forth between data collection and analysis. For example, needing to plan additional data collection or adaptation of interview questions in light of the data already collected or modifying the interview questions from one interview to the next (Dicicco-Bloom & Crabtree, 2006; Ramani & Mann, 2016). This concurrent process of data collection and analysis also assists researchers in determining when data saturation and theoretical sufficiency have been reached (Schut et al., 2018).

In qualitative research, particularly when analysing transcripts, software programmes are generally used to, “Facilitate management of large data sets, organisation of data into codes and coding categories, linking codes and memos and converting it into a searchable database” (Ramani & Mann, 2016). In this study, NVivo for Mac software was used for the analysis of said interview transcripts.

4.3.6.1. Phenomenographic analysis

In line with the Phenomenographic research methodology selected for the sub-study on conceptions, Phenomenographic analysis was used for uncovering lecturers’ varied conceptions of assessment.

In Phenomenographic analysis it is often said that conceptions, the categories of description, “emerge” from the data, but this is not a passive process, rather, it requires the active role of the researcher, in that the researcher is the one who is reading, interpreting, making decisions and drawing out these themes, with reflexivity being practiced throughout (Åkerlind, 2005; Braun & Clarke, 2006; Ramani & Mann, 2016). One author suggests that the term “immersion” be used as opposed to “emergence” to demonstrate the activity in this process (Varpio et al., 2017). Regardless of whether conceptions are “discovered” or “constructed”, analysis must remain faithful to the data, honestly reflecting and representing the participants’ understandings, with justifications for conceptions provided through quotations given, indicating interpretative awareness and the trustworthiness of said findings (Bowden & Walsh, 2000; Stenfors-Hayes et al., 2013).

Phenomenographic analysis traditionally consists of seven steps: familiarisation, complication, comparison, grouping, articulation, labelling and contrasting (see **Figure 4.3.**) (Bowden & Walsh, 2000; Gillsjö et al., 2011; Holland et al., 2013; Marton & Pong, 2005; Moller et al., 2010; Munck et al., 2012a, 2012b; Pihl et al., 2011; Sandy, 2014; Sjostrom & Dahlgren, 2002; Stenfors-Hayes et

al., 2011, 2012; Wahlstrom et al., 2001; Wilhelmsson et al., 2011). (1) Familiarisation ("immersion") occurs through transcription of audio-recorded interviews and multiple subsequent readings of these transcripts: transcription is a transformational process (for audio is transformed, translated, into text), and transcripts are read multiple times, ensuring checking, probing and testing of findings (Bowden & Walsh, 2000; Dortins, 2002). It is suggested that more than a single researcher reviews the transcripts ("analytical collaboration") to ensure that the data speaks for itself and that the results are not heavily dependent on the professional interpretation and judgement of a single researcher (Tight, 2016). While the majority of data collection and analysis was performed by a single researcher in this study (the author of this thesis), additional independent researchers were consulted (independently reading and coding transcripts), followed by rounds of rich discussion, until a collective consensus was achieved.

The second step of Phenomenographic analysis, (2) compilation, includes searching for statements that correspond to the aim of the study, which form the initial codes: phrases of interest are marked or highlighted, with the meaning of the phrases being determined in two phases, firstly, through considering the phrase within the context of its transcript as a whole, and secondly, all the phrases together ("pool of meanings") are considered collectively (separate from their individual transcripts and across all transcripts) (Bowden & Walsh, 2000; Stenfors-Hayes et al., 2013; Tight, 2016; Woollacott et al., 2013). These phrases act as evidence for a conception and eventual categories of description.

During the (3) comparison step, similarities and differences are noted. Then (4) grouping examines the phrases highlighted and brings similar concepts together, which forms tentative categories of description that reflect the emerging conceptions (Bowden & Walsh, 2000). During the grouping step, all the phrases are grouped together or separated (Bowden & Walsh, 2000). During categorisation, proposed categories are debated, rearranged and altered until data saturation, theoretical sufficiency and consensus are reached. These categories are then organised into a logical structure (hierarchy), that reflects the natural relationship between categories, "With each higher level encompassing those below it, and the highest level

representing the most advanced or developed way of experiencing the phenomena" (Bowden & Walsh, 2000; Tight, 2016). These categories represent different ways (key meanings) of understanding something. They are related (not independent), and collective (not individual) (Ojo & Booth, 2009; Stenfors-Hayes et al., 2011; Tight, 2016; Woollacott et al., 2013). (5) Articulation then re-examines the codes and refines the categories further if needed. This is then followed by (6) labelling these groups of conceptions, into categories of description, to reflect their meanings. Lastly, (7) contrasting takes place, in the form of comparison between similar and different categories of description. Details on how each dimension and category of description in the Outcome Space came to be will be discussed in the results section below (**Chapter's 5** and **6**).

This 7-step process requires multiple iterations, with constant interplay, to reveal gaps, uncertainties, new emergences and distinctions, until greater clarity is achieved, with a decreasing rate of change, until the categories of description possess a stable structure (Bowden & Walsh, 2000; Stenfors-Hayes et al., 2011; Woollacott et al., 2013). If more than one researcher is involved in data collection and analysis, then consensus is negotiated through deep discussion until full agreement is reached by all researchers, which took place in this study (Gillsjö et al., 2011; Wahlstrom et al., 2001).

The end product of Phenomenographic analysis, or the sum of the resultant categories of description, is the Outcome Space (Åkerlind, 2005; Ashworth & Lucas, 1998; Larsson & Holmström, 2007; Marton, 1981; Marton & Pong, 2005; Richardson, 1999; Stenfors-Hayes et al., 2013). The Outcome Space consists of categories of description, which are predominant patterns that have emerged from the data and represent the collective ways of conceptualising (experiencing and understanding) a phenomenon, and the dimensions (critical features) that make up a conception (Wilhelmsson et al., 2011). The inclusion of various dimensions within conceptions allows for a generous description of a phenomenon (Holland et al., 2013; Ojo & Booth, 2009; Pihl et al., 2011; Wilhelmsson et al., 2011; Woollacott et al., 2013). The categories of descriptions title the *variations* in conceptions of experiences of the phenomenon and *not*

common conceptions of a phenomenon, for, they are not about the phenomenon itself, but, the variations of people's experiences of the phenomenon (Cibangu & Hepworth, 2016; Tight, 2016). The criteria for each category of description are that each category must be distinct from another and logically related to the other categories, as is displayed in the relational organisation of the Outcome Space (Åkerlind, 2005; Harris, 2011; Larsson & Holmström, 2007; Marton, 1981; Richardson, 1999; Stenfors-Hayes et al., 2012; Stenfors-Hayes et al., 2013).

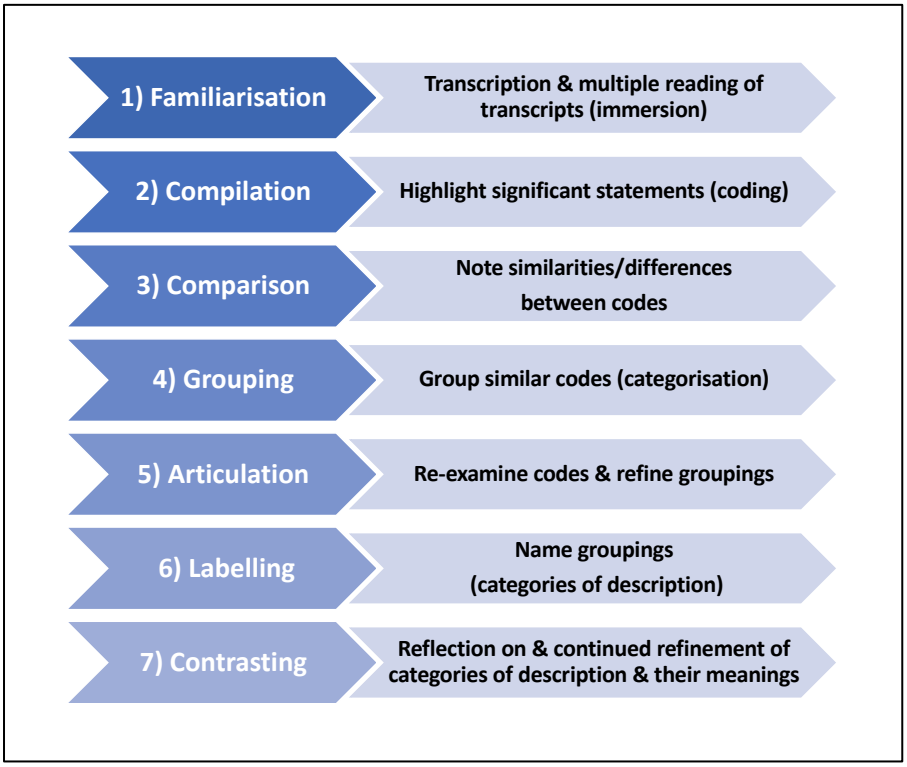


Figure 4.3.: Phenomenographic analysis: an overview. Phenomenographic analysis of collected interview data consists of 7 steps: (1) Familiarisation, (2) Compilation, (3) Comparison, (4) Grouping, (5) Articulation, (6) Labelling and (7) Contrasting, resulting in an Outcome Space describing a phenomenon as various categories of description according to a number of dimensions (Gillsjö et al., 2011; Holland et al., 2013; Moller et al., 2010; Munck et al., 2012a, 2012b; Pihl et al., 2011; Sandy, 2014; Sjostrom & Dahlgren, 2002; Stenfors-Hayes et al., 2011, 2012; Wahlstrom et al., 2001; Wilhelmsson et al., 2011).

The Outcome Space is usually presented in a two-dimensional depiction, such as a table, with the categories of description, each representing a conception, along the top, and the dimensions along the side. The Outcome Space describes and depicts the relationship between said

categories and their dimensions (see **Figure 4.4.**) (Gillsjö et al., 2011; Wilhelmsson et al., 2011; Woollacott et al., 2013). The main criteria of a Phenomenographic Outcome Space are that, "Each category reveals something distinctive about a way of understanding the phenomenon, the categories are logically related, and the critical variation seen in the data is represented by as few categories as possible" and that the categories are defensible (supported by quotations), useful and meaningful (provide value to the intended audience) (Stenfors-Hayes et al., 2013).

	Categories of description		
	Conception 1	**Conception 2**	**Conception 3**
Dimension 1			
Dimension 2			
Dimension 3			

Figure 4.4.: Phenomenographic analysis outcome space: an example. An outcome space consists of various categories of description of the main groupings that have arisen and their accompanying dimensions or the specific and detailed characterisations and variations within those categories. These categories of description represent the conceptions, along with the elements they contain (Åkerlind, 2005; Ashworth & Lucas, 1998; Larsson & Holmström, 2007; Marton, 1981; Richardson, 1999; Stenfors-Hayes et al., 2013).

The phenomenographic analysis described thus far was utilised to answer the research question of the sub-study, relating to lecturers' conceptions of assessment. However, as this study also sought to explore lecturer assessment practice, and in line with HBT (Cilliers et al., 2015; Glanz et al., 2008), what personal and contextual factors impact on their assessment practice, interviews were conducted in two parts. The first part of the interviews explored conceptions, while the second part investigated factors. thematic analysis then performed on the second part of the data collected in order to answer the research question on factors influencing lecturer assessment practice.

4.3.6.2. Thematic Analysis

thematic analysis, a form of content-based analysis, is an accessible and flexible approach to qualitative data analysis, that may be used across a range of theoretical and epistemological positions, and is compatible with diverse paradigms (Braun & Clarke, 2006, 2012). In fact, Clarke and Braun (2013) describe thematic analysis as, "Just an analytic method, rather than a methodology." Thematic analysis may be defined as, "A method for systematically identifying, organizing, and offering insight into patterns of meaning (themes) across a data set" (Braun & Clarke, 2012). "Themes" refers to implicit and explicit ideas within the data (Guest, MacQueen, & Namey, 2012). As thematic analysis allows researchers to make meaning across multiple data sets, it may be used to make sense of collective experiences (Braun & Clarke, 2012).

Thematic analysis may be inductive (data-driven, a "bottom-up" approach) or deductive (theory- and hypothesis-driven, a "top-down" approach, in that a framework is imposed or applied to the data) or both (Braun & Clarke, 2006, 2012; Fereday, 2006). In this study thematic analysis was both inductive, allowing the data to speak for itself with no prior expectations as themes emerged, and deductive, using the general model of HBT as a sensitizing concept (for example, using "personal" and "contextual" as broad themes for coding and interpretation) (Braun & Clarke, 2012; Fereday, 2006). Importantly, the deductive categories are not to confine data analysis but, rather, to guide it (Fereday, 2006), and, in reality, "It is impossible to be purely inductive" (Braun & Clarke, 2012).

The process of thematic analysis consists of 5 steps (see **Figure 4.5.**), starting with (1) familiarizing oneself with the data, which is also referred to as "immersion": for example, through listening to the recorded interviews, transcribing the recorded interviews and reading through the transcribed texts (Braun & Clarke, 2006; Fereday, 2006). Next (2) initial codes (labels) are generated through identifying important pieces of information prior to the process of interpretation (Braun & Clarke, 2006), capturing some of the, "Qualitative richness of the phenomenon" (Fereday, 2006). Codes refer to features of the data that appear to be of interest

to the researcher, importantly, coded data differs from units of analysis, as coded data is grouped into broader themes (the units of analysis) (Braun & Clarke, 2006). These codes may be developed through the familiarisation of the data (an inductive approach) or guided by the research question or theoretical framework of the study (a deductive approach) or a combination of both. As the data is coded, phrases of transcribed interview text are highlighted and ascribed to certain codes. These phrases describe the meaning underpinning the theme or provide an explanation of/for the code (Fereday, 2006). All the relevant data should relate to the codes, patterns and themes identified, for, if new or different theme emerges, then these initial codes must be expounded upon (Aronson, 1994; Braun & Clarke, 2006; Fereday, 2006). Coding must also be performed inclusively, mindful of its surrounding context (Braun & Clarke, 2006). By the end of the coding phase, text may either be uncoded, coded once or coded multiple times under different codes (Braun & Clarke, 2006).

The next step is (3) searching for themes from the coded data, refocusing the analysis at a broader level. This has been described as "coding your codes" or looking for similarity, coherence and meaningfulness in the data; coded data is now collated to relevant, over-arching themes (Clarke & Braun, 2013). This collation of themes brings together fragmented ideas or experiences that lack meaning alone, for the gathering of themes presents a comprehensive view of the data (Aronson, 1994; Braun & Clarke, 2006). Reviewing the themes (4) follows to ensure that the themes tell a convincing, compelling and coherent story of the data: do these themes accurately reflect the data as a whole? This means that themes may be expanded, split, discarded or redeveloped (Clarke & Braun, 2013). The last step in the thematic analysis process is that the final themes need to be (5) defined and named through corroborating and legitimating of coded themes; there must be an interaction between the data, codes and themes, consistency and evidence for the final coded themes (Braun & Clarke, 2006; Fereday, 2006). Evidence for said themes, in the form of quotations from the research interviews, are provided in the results chapters.

While there appears to be overlapping steps in the analytical processes of thematic and Phenomenographic analysis, a characteristic of Phenomenography is the hierarchical and relational nature of the Outcome Space, with categories and dimensions interacting with one another to, together, describe a conception (Stenfors-Hayes et al., 2013). This relational aspect is not present in thematic analysis (Stenfors-Hayes et al., 2013). However, as the themes identified relate to potential factors in the creation of a model of behaviour, ultimately, these factors are relational, for HBT posits that personal and contextual factors interact in order to direct behaviour.

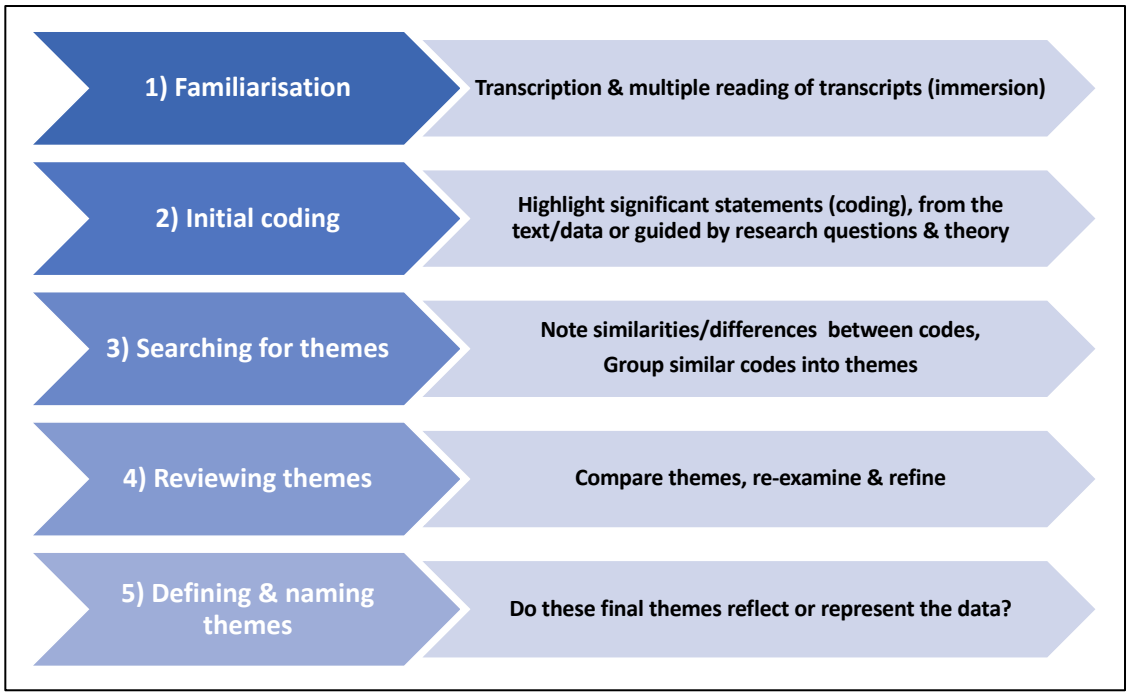


Figure 4.5.: Thematic analysis: an overview. Thematic analysis consists of 5 steps: (1) familiarisation, (2) initial coding, (3) searching for themes, (4) reviewing themes and (5) defining and naming themes. It must be remembered that, although there is a step-by-step approach to thematic analysis, it too is an iterative and reflexive process (Fereday, 2006). Adapted from Braun and Clarke (2006).

To summarise, purposive and maximum variant sampling strategies were employed to select for individual final year/s course convenors from medical HEIs, in various disciplines, from diverse Southern settings. Data collection took place in three stages: initially in South Africa before the findings were further clarified, refined and validated in additional Southern contexts. In order to

deeply explore lecturers' conceptions of assessment and the factors influencing their practice, semi-structured interviews took place, either in person or online, with recorded interviews being transcribed verbatim for following Phenomenographic and thematic analyses, in a parallel analysis process, to determine lecturers' conceptions and factors influencing their assessment practice respectively.

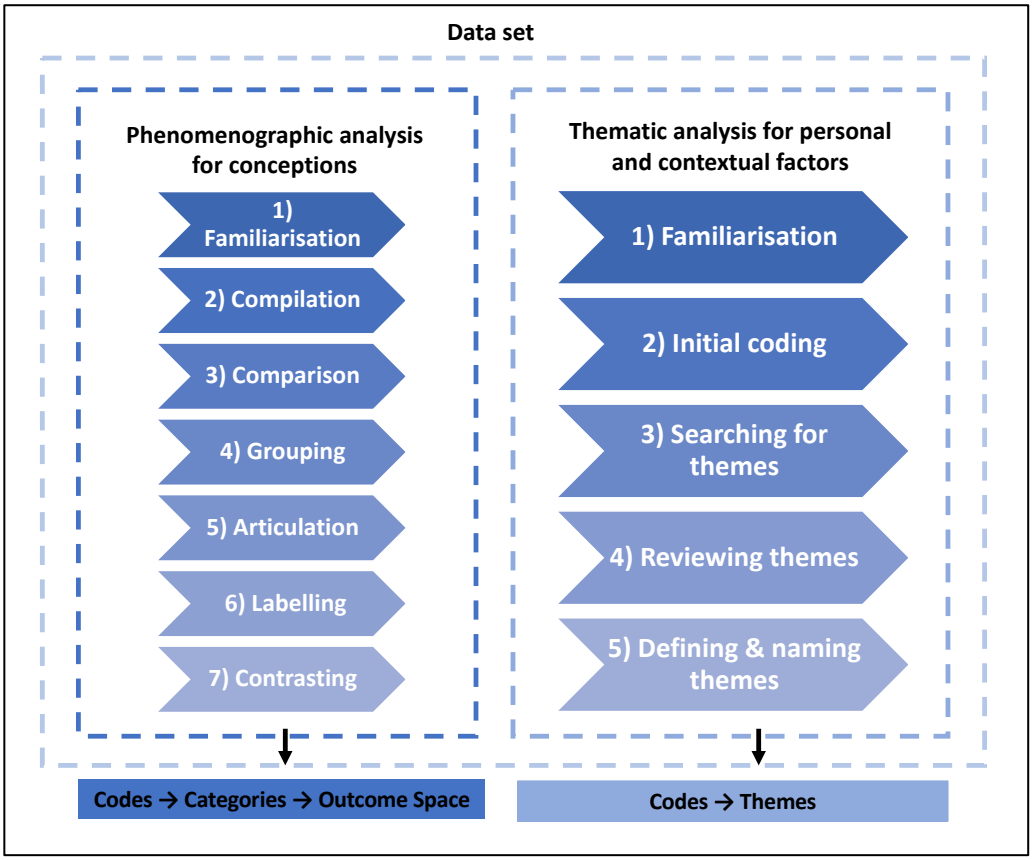


Figure 4.6.: Parallel data analysis process. Phenomenographic analysis, focused on conceptions of assessment, and thematic analysis, focused on personal and contextual factors influencing assessment practice, were conducted in a parallel process.

The findings from the sub-study on conceptions is presented first, and the factors identified presented second, for the HBT framework is "an embedded reciprocal system" for organising the relationship between factors at various levels that impact on behaviour. I am arguing for conceptions at the centre of the concentric circles, followed by additional personal, proximal and distal contextual factors. Throughout the entire research process, from sampling to data

collection to analysis, principles of reflexivity, rigour and ethical research were practiced to ensure both high quality, credible and socially just research outcomes.

Chapter 5: Pilot study: Lecturers' Conceptions and Practice of assessment in a South African context

The aims and objectives of this study were to explore lecturers' conceptions of assessment in a range of diverse, Southern contexts (**Chapters 5** and **6**), as well as what personal and contextual factors influenced their assessment practices (**Chapter 7**). To this end, in total, 31 semi-structured interviews took place, amounting to 231 000 transcribed words: 12 interviews were conducted at an HEI in South Africa (referred to as the first phase or "pilot study" and described here in **Chapter 5**), at a second HEI in South Africa and at an HEI in Mexico, accounting for 6 and 13 interviews respectively (referred to as the second and third phases of the "confirmatory study" and described in **Chapter 6**). The possible link between espoused (held) and enacted conceptions (practical manifestations of lecturer assessment behaviour) are described. Investigation of further personal and contextual factors influencing lecturer assessment practice are described in **Chapter 7**. A final discussion of these findings, in comparison to current literature, and in light of our conceptual frameworks of HBT and Southern Theory, are found in **Chapter 8**.

5.1. Phenomenographic Methodology: Outcome Space organisation

Throughout the course of this research project, three Outcome Spaces were developed. The first Outcome Space reflected the first phase of the study, the pilot study, based on the first South African dataset alone (transcripts "SA1-12"). This Outcome Space was developed through inductive Phenomenographic analysis, before beginning data collection at additional Southern sites, which was referred to as the confirmatory study. The second Outcome Space reflected the second phase of the study. This Outcome Space was also developed using inductive Phenomenographic analysis, considering the second South African data set alone (transcripts "SA13-18"), before comparing it to the first Outcome Space. The third Outcome Space was the final, collective Outcome Space, and it represented *all* the data collected. This Outcome Space was the result of both inductive and deductive Phenomenographic analysis. First the data

collected from Mexico, the third phase of this study, was analysed inductively (transcripts "MX1-13"), with an open mind and without preconceived ideas or agendas, allowing the data to speak for itself, leading to the outline of a preliminary Outcome Space. Deductive analysis, in terms of comparing the third Outcome Space to the previous two, rereading all the transcripts and utilising relevant literature for sensitizing concepts, led to the development of the final Outcome Space, representing a coherent result from all the data collected from three diverse Southern settings.

The three Outcome Spaces, and their categories of lecturers' conceptions of assessment, were titled differently. The conceptions from the **pilot study** were titled "**Undirected**", "**Content-focused/Reproduction-directed**" and "**Competency and Conceptually-focused/Application-directed**" (see **Table 5.1.**). The conceptions from the **second phase** of the study were titled "**Detached practitioner**", "**Emerging equilibrium**" and "**Engaged educator**" (see **Table 6.1.**). The conceptions of the **third phase** and **final Outcome Space** (representing *all* the data collected, data saturation and theoretical sufficiency) were titled "**Passive operator**", "**Awakening inquirer**", "**Active owner**" and "**Scholar**" (see **Table 6.4.**). While, at first glance, these Outcome Spaces may appear quite different, and incongruous, I argue that they reflect the deepening of the processes of analysis, abstraction and theorisation, as well as the maturation of the researcher over time (Morse, 1994). It could be said that the differences between the three Outcome Spaces are due to a qualitative difference in the researcher, as opposed to a qualitative difference between the different data collection contexts. Further discussion of each, and all, Outcome Spaces are detailed in **Chapter 8**.

In considering the organisation of Outcome Spaces, in general, while an Outcome Space may appear to have discrete dimensions and categories, as represented in a two-dimensional table, in reality, the Outcome Space is more *fluid*. It represents a *spectrum* or *continuum*, the diversity and range, of conceptions, with dimensions of that conception relating to and connecting with other conceptions. The lines between categories are somewhat blurred. Lecturers may find that a category of a conception resonates strongly with them, however, they may also hold a mixture of conceptions or possess varied dimensions of more than a single conception.

The Outcome Space also represents the *collective* way a group of lecturers view and understand assessment. Individual lecturers may not fit neatly or exclusively into a single category, but across multiple categories. Indeed, "Categories should be as faithful as possible to the individual's conceptions, but are not equal to them as the conceptions are dynamic and represent a relationship between the individual and his or her context" (Stenfors-Hayes et al., 2013).

Phenomenographic Outcome Spaces are relational and hierarchical. This means that the first conception is usually the simplest, with subsequent categories becoming more advanced or complex as they progress, encompassing and surpassing previous categories (see **Figure 5.1.** as an illustration of this concept). For example, the Competency and Conceptually-focused/Application-directed conception of assessment incorporates the Content-focused/Reproduction-directed conception of assessment, along with new and more nuanced variations of dimensions.

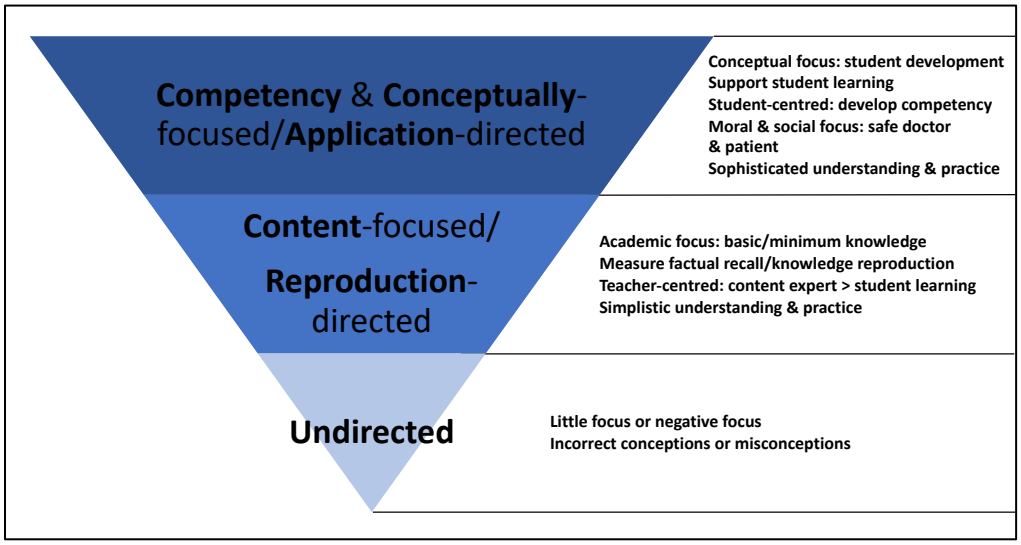


Figure 5.1.: The hierarchical nature of the Phenomenographic Outcome Space: Pilot study example. The first category or conception is the "simplest" with the final category or conception representing the most "sophisticated" level of conception, therefore each level encompasses and builds on and thereby surpasses the previous level. Each category or conception thus becomes more complex and advanced than the previous category, displaying the hierarchical nature of the Outcome Space.

5.2. Pilot Study: Lecturers' Conceptions of Assessment in a South African context

In order to deeply explore lecturers conceptions of assessment, a pilot study was performed in which 12 semi-structured interviews with final year/s medical course convenors, from a range of disciplines, took place at a South African HEI (traditionally Afrikaans in language and culture), totalling approximately 127 000 transcribed words.

Phenomenographic analysis was performed on the transcripts and three categories of description of assessment, along with three dimensions, were found. The three conceptions were labelled "**Undirected**", "**Content-focused/Reproduction-directed**" and "**Competency and Conceptually-focused/Application-directed**", and the three dimensions were called "purpose", "learning effect" and "technical knowledge" (see **Table 5.1.** and **Appendix: Table 5.1.** for a table with illustrative quotes). Theses dimensions referred to what participants viewed to be the function ("purpose") of assessment, what learning outcome ("learning effect") they believed assessment achieved, and their awareness and understanding ("technical knowledge") of key principles of high-quality assessment practice.

Table 5.1.: Pilot study: An outline of the Phenomenographic Outcome Space reflecting one group of South African lecturers' conceptions of assessment.

Outcome Space	Undirected	Content-focused/ Reproduction-directed	Competency and Conceptually-focused/ Application-directed
Purpose	• None or negative	• Summative/psychometric measurement • Formative/learning • Teacher-centred	• Formative: student-centred • Social and moral: patient and society-centred
Learning effect	• None or negative	• Directs superficial learning • Short-term effect	• Directs deeper understanding • Long-term effect
Technical knowledge	• None or incorrect	• Simple and developing	• Sophisticated understanding

5.2.1. Pilot Study: The Dimensions of Lecturers' Conceptions of Assessment

The first phase of this study resulted in an Outcome Space with three categories of description, consisting of three dimensions for each conception. In hindsight, the dimensions were prescriptively shaped by the interview questions which explored the same three topics: the purpose of assessment, the effect assessment has on learning, and their awareness and understanding of key criteria and principles for designing, practicing and achieving high quality assessment.

For example, looking at the dimension of the "purpose" of assessment, participants were explicitly asked their thoughts on the function of assessment, what they believed it accomplished and for what reason they practiced it. Similarly, for "learning effect", participants were asked if they thought that assessment drove learning, and, if so, in what way. For "technical knowledge" questions centered around what participants thought were the requirements or components of a "good" assessment practice. They were asked about how they designed their assessments, what informed their decision-making and so on. The varied responses from participants formed the range of each dimension.

The titles of each category of description, reflecting lecturers' varied conceptions of assessment, were labelled according to what was believed to best represent the entire category, and its dimensions, that had logically sorted together, within individuals, and across transcripts. I have also included, beneath each quotation (in brackets), some "tags", similar to the coding used during analysis, to assist the reader in understanding my thinking and reasoning behind the inclusion of each quote as evidence of and support for the statements I have made.

5.2.2. The Undirected conception of assessment

The **Undirected** conception of assessment was held by participants who had little time and ambition for their assessment practice. These participants were "undirected" in their thoughts and understanding of assessment. They believed that assessment served little to no valuable purpose and negatively referred to it an "irritation" (SA3) or "necessary evil" (SA6).

> "So, students come through, we just assess them and throw them out" (SA8).
> (Little useful purpose, negative perception)

> "The path to hell is paved with good intentions. So, I think it's *a necessary evil*" (SA6) – *emphasis* added.
> (Negative perception)

In line with the perception that assessment serves little meaningful purpose, participants holding Undirected conceptions were not convinced that assessment drives student learning. Some used assessment as a negative motivator of learning, focusing on the negative emotions associated with assessment (stress, anxiety, fear), over the idea of using assessment to support student learning in a positive manner.

> "How well students respond to the *threat of a test*" (SA6) – *emphasis* added.
> (Negative driver of learning, punitive)

> "For me the Holy Grail would be to have a system where evaluation is a process of inspiration and motivation and *just a bit of a fright*… Respect is good, but *fear is better*, you know? So just *a bit of a scare*… so the student will have the drive to know the work… To get their [students] *stress levels up*, because I also believe that without that bit of *adrenaline pumping* you don't really learn… I want there to be a bit of *scare tactics*… because you're not going to study if you're not the tiniest bit *anxious*… It's this whole thing about them *being scared enough that they actually sit down and study*… An

examination … is *never a one hundred percent pleasurable experience*" (SA1) – *emphasis* added.

(Negative driver of learning, punitive)

When it came to the technical knowledge, or understanding of the principles behind quality assessment practice, participants who held an Undirected conception gave little thought to what it meant to design and deliver a "good" (valid and reliable) assessment. These participants also held simplistic or incorrect ideas of assessment principles. For example, when it came to determining the cut-score (pass mark) for an assessment, these participants openly admitted that it was an arbitrary "thumb suck" for them, or a blind following of what was suggested by HEI rules, rather than a process guided by standard setting.

"I don't think there is any clear objective way and it's done utterly on *thumb suck*" (SA6) – *emphasis* added.

(No or incorrect technical knowledge, subjective)

"It's [marking] just sort of *gut feel* … A *thumb suck*" (SA12) – *emphasis* added.

(No or incorrect technical knowledge, subjective)

"It's [marking] a very *subjective* thing … I have to say it's *a gut feeling …* There is no objective or quantitative measuring instrument I use" (SA3) – *emphasis* added.

(No or incorrect technical knowledge, subjective)

"Is 50 the right mark? To this day I don't know why it's said that 50 is the pass mark. Whether it's scientifically proven or whether 50 sounds better than 52 or 49 – I'm not so sure about it" (SA10).

(No or incorrect technical knowledge)

"The *assumption* that the average mark would be somewhere around about sixty-five, and of course that's *a complete assumption*. If you see marks that are substantially less than that then *you can adjust those or not as you see fit*" (SA6) – *emphasis* added.

(No or incorrect technical knowledge, subjective)

"I think it's [50%] an arbitrarily chosen figure" (SA11).

(No or incorrect technical knowledge)

"There was an unacceptably high failure rate and there was something like sixteen people who got less than forty percent, which I thought was a very poor reflection on the team of lecturers. I discussed it with the person who was then the responsible party … and *it was decided to make an adjustment* [to the marks]" (SA3) – *emphasis* added.

(No or incorrect technical knowledge, subjective)

"We have years of *tradition* where if someone has 74%, we automatically make it 75%" (SA10).

(No or incorrect technical knowledge, follow tradition)

"Whether it be seventy-five, seventy-seven, or seventy-three, it is all the same to me. I think it's kind of a *historical thing* that, that it [a distinction] came to centre around seventy-five … I think the mark … is probably just *a number of comfort*" (SA2) – *emphasis* added.

(No or incorrect technical knowledge, follow tradition)

To summarise the **Undirected** conception of assessment, assessment was conceptually understood to serve no to little meaningful purpose and viewed negatively, an irritation and burden, and was used negatively to motivate of student learning. Participants who held this conception also possessed incorrect to poor technical knowledge of assessment, giving little critical thought to the design and practice of their assessment, instead, often relying on external rules and tradition over evidence-based practice.

5.2.3. The Content-focused/Reproduction-directed conception of assessment

Participants with a **Content-focused/Reproduction-directed** conception of assessment viewed assessment as serving multiple purposes and emphasized a more content-focused and reproduction-directed aim and practice. While participants may have viewed assessment as formative, a tool to drive student learning, and as a gatekeeping mechanism, in order to protect the standards and reputation of the profession and HEI, it was mostly viewed as a summative tool to ensure that the content basics were covered and reproduced by the students, as seen through achieving a pass mark. This displayed a teacher-centred approach to assessment, as these participants saw themselves as content-experts whose role it was to impart information to students. While assessment was used to motivate learning, the desire was a more superficial reproducing of the minimum factual information than student development. This differed from the Undirected conception of assessment and negative motivator of assessment, for, while the Content-focused/Reproduction-directed conception of assessment also led to negative learning outcomes, in the form of memorization and short-term learning, there was no overt intention to cause negative emotions in students as a way of driving studying.

> "I know assessment is a very important motivator to get people studying" (SA7).
> (Assessment drives learning)

> "[Testing the] minimum level of factual knowledge" (SA3).
> (Content reproduction, summative, superficial learning, short-term)

> "[That students have] grasped the basic ideas" (SA12).
> (Content reproduction)

> "I'm going to ask you this, therefore you must learn it" (SA6).
> (Content reproduction, teacher-centred, superficial learning, short-term)

> "To ensure that the whole class pays attention to … what *the lecturer thinks is important*" (SA12) – *emphasis* added.
>
> (Teacher-centred)

Related to the idea of teacher-centredness, these participants saw themselves as experts and, therefore, gatekeepers of their profession. They viewed assessment as a gatekeeping device to control who entered into their profession. These participants firstly felt a responsibility towards their profession and HEI, and secondly to the public.

> "There has to be an … almost uniform standard against which you have to assess people… I am not the one who holds the key, it's not for me to decide when the door is unlocked and when it is locked [but] I am a representative of the profession … I can't allow someone to go into the community, to sign off… and we let him go and then he kills people or he makes terrible errors in judgement. He should never have … I mean that's accusation against us. Now, most of my colleagues feel that we have a tremendous responsibility towards the public and the university and our profession … We have to be unprejudiced. But there has to be some level of … expertise and knowledge and skills where we can say look this person isn't suitable for the job market" (SA7).
>
> (Disciplinary knowledge and profession-centred, gatekeeping)

> "If he [a student] doesn't succeed … To be fair to the other students and to the community, we can't allow a guy … like that to pass through the gate" (SA11).
>
> (Disciplinary knowledge and profession-centred, gatekeeping)

The focus of the Content-focused/Reproduction-directed conception of assessment was to ensure that students were able to reproduce the content they had been taught. An unfortunate learning effect of this focus was a superficial, short-term learning, whereby participants believed students merely memorised and promptly forgot what they had learned once the assessment was done. This was also reinforced by their summative view of assessment where student learning was represented by an assigned score, which led students to be more mark-focused,

instead of considering the longer-term perspective of developing competency for future clinical practice and patient care.

> "At the moment I am very sorry to say that it [assessment] only stimulates them [students] to go and sit and cram the content two days before the time … [assessment] leads to … *very little long-term retention*. The student is studying something that he saw two weeks ago and then [not] to remember it until now … and then he'll repeat the process at the end of the year when he writes his block [end of year] examination … that's not how one learns … *it's counter-productive*" (SA1) – *emphasis* added.
> (Content-reproduction, superficial learning, short-term)

> "The assessment … is purely for a *short-term goal* and has *no real meaning* for the student in terms of the following year" (SA6) – *emphasis.*
> (Content-reproduction, superficial learning, short-term)

> "What happens then is that the students gear themselves towards acquiring *short-term knowledge and short-term recall*, and when they walk out of that exam room, all that gets flung out by the window … I'm not sure whether exams are then a tool by which you can measure whether you've equipped the student appropriately" (SA11).
> (Content-reproduction, superficial learning, short-term)

In terms of the dimension of <u>technical knowledge</u>, participants with a Content-focused/Reproduction-directed conception of assessment displayed a more sophisticated knowledge of assessment than participants with an Undirected conception. This developing knowledge was seen in their ability to recognise a problem or acknowledge limitations in their assessment design and practice. However, these participants were unable to provide a correct solution or implement a high quality assessment practice themselves, for they did not yet possess the highest evidence-based dimension of the Competency/Conceptually-focused/Application-directed conception of assessment. For example, a lecturer may have desired a more "objective"

type of assessment, yet, expressed their uncertainty in how they may be able to do so. This was also displayed in their lack of confidence in their assessment practice.

> "I hope we don't make mistakes that often, but I have a sinking feeling that we do make them quite often. *Yes, we make mistakes*, as you say, because we base our decisions on instruments of assessment that are probably *not as accurate as we hoped* … We question the *validity* of it [our assessment practice]" (SA2) – *emphasis* added.
> (Questioning/uncertain, developing assessment knowledge)

> "*I don't think the assessment tool may be that accurate to differentiate at that level* [the difference between a student achieving 48% versus 52%], because it may be the question of just one multiple-choice question that made the difference for those students. So … you'd probably need *a wider spectrum*, rather than a final point, you know, in terms of satisfactory and unsatisfactory" (SA11) – *emphasis* added.
> (Questioning/uncertain, developing assessment knowledge)

> "It doesn't tell me the person is, you know has seventy-five percent of the knowledge they're supposed to know, that is required of him in this subject field. *I don't think our instruments of assessment are sensitive enough* for me to do that, that we can say, "If someone gets seventy-five percent we can be sure that he has seventy-five percent." That's impossible, *we don't have such sensitive instruments by which we assess*, that's all" (SA2) – *emphasis* added
> (Questioning/uncertain, developing assessment knowledge)

To summarise the **Content-focused/Reproduction-directed** conception of assessment, participants conceptually understood the purpose of assessment as both summative and formative, summatively to measure that students reproduced the minimum required level of factual knowledge, and formatively, as a tool that motivated student learning (memorisation, in this instance). Unfortunately, these participants also saw the short-term outcome and mark-focus of their students. In terms of their assessment knowledge, these participants held emerging conceptions of criteria for quality assessment practice, as was displayed in the expressed confusion or uncertainty about the technical aspects of their assessment. However, this revealed a basic awareness of concepts such as validity and reliability. It could be proposed that this awareness was based, at least partially, on tacit craft knowledge and expertise gained from practical assessment experience over time, even if, at this stage, they were unable to fully articulate or implement it in their assessment practice.

5.2.4. The Competency and Conceptually-focused/Application-directed conception of assessment

Participants with a **Competency and Conceptually-focused/Application-directed** conception of assessment understood assessment to have multiple purposes, but emphasised the formative, moral and social functions. Formatively, they saw assessment as a tool that contributed towards student development. These participants also viewed assessment as a moral and social practice that ensured competency for future clinical practice.

> "There's no question that at the end of the day, we're *responsible to civil society*. So, we train the students to assimilate their knowledge and technical skills that would make them *good practitioners*" (SA8) – *emphasis* added.
>
> (Moral and social purpose, competency, long-term)

> "In the long term it [assessment] has to do with … the *certifying someone as competent* to a particular task – *the responsibility towards society*" (SA12) – *emphasis* added.
>
> (Moral and social purpose, competency, long-term)

As with Content-focused/Reproduction-directed conception, the dimensions of assessment purpose and assessment's learning effects were linked. In this instance, Competency and Conceptually-focused/Application-directed participants viewed assessment's learning effects as a formative support in developing academic, social and moral competencies (as illustrated in the quotes above). Students, therefore, needed to move beyond superficially memorising knowledge for reproduction to deeply understanding and applying said knowledge, particularly in clinical settings, as, ultimately, the goal was to produce competent practitioners that served society and treated patients safely. This view, to use assessment to drive student learning, growth and development, was displayed in the asking of clinically-based questions that tested students' insight, and the provision of feedback, which reflected a more advanced understanding, knowledge and practice of assessment than previous categories.

> "[The] *application* of factual knowledge" (SA3).
>
> (Application of knowledge)

> "Assessment, in the end, has to be a reflection of … how the student has grown in terms of that subject field; how he has *developed* … *Higher order* questions … *insight* … I think the more *clinical* the questions are, the more scenario-based ... case studies … It forces them [students] to integrate the work … A *clinical* problem scenario … a *patient* gets sketched for them every time. They are brought to a virtual space which reflects how their *practice* might look someday … important skills that a *doctor* has to have … [Assessment] has to be in line to help students *improve* their skills or their abilities to become *better doctors*" (SA4) – *emphasis* added.
>
> (Assessment drives learning, student-centred, deep understanding, long-term)

"Assessment is a great mechanism for giving *feedback* to students so that they will know where they stand … and what effort is needed to get to where they need to be" (SA2) – *emphasis* added.

(Assessment drives learning, formative, student-centred, sophisticated assessment knowledge)

"To offer *feedback* to a student about his knowledge and to indicate where the defects lie … Actually, the ideal is that the students can be *reflective* about their assessment" (SA4) – *emphasis* added.

(Assessment drives learning, formative, student-centred, sophisticated assessment knowledge)

"I think there is not enough *discussion* with individual students about their marks … I think there might be a way of actually being able to *encourage* students to work to their … more optimal *potential* … [And] it shouldn't just be the weaker students that are spoken to … there is someone *listening on a regular basis*, *someone taking an interest*, *right from the beginning*" (SA12) – *emphasis* added.

(Assessment drives learning, formative, student-centred, sophisticated assessment knowledge)

In contrast to the negative motivation of the Undirected conception of assessment, participants in this category held positive views of assessment as a motivator for driving desirable study behaviours and learning outcomes in students, which was also in contrast to the superficial, short-term learning effects of the Content-focused/Reproduction-directed assessment conception.

"They [students] fared very poorly in the first test … I've talked to the few who've come to see me, and I sort of told them, "But you guys *don't study correctly*, you don't use *insight* … if you use *comprehension*, you'd be able to answer the factual question better … you'd learn more effectively if you'd learned at a *deeper level*, even for answering the superficial questions"… I preached to a lot of them … What's interesting is that I was so impressed with … how well they fared in that question … a great percentage fared very

well … Even one test could easily influence the students' behaviour. They'd already realised that *their strategy* wasn't working" (SA4) – *emphasis* added.

(Assessment drives learning, formative, student-centred, deeper understanding)

In terms of the dimension of technical knowledge for a Competency and Conceptually-focused/Application-directed conception of assessment, participants held conceptions of reliability, validity, educational impact (displayed in the learning effects mentioned above), practicality and acceptability. These participants expressed views that assessment should be representative and cover the entire curriculum (validity), students should have multiple assessment opportunities (reliability and validity), assessment should be clinically-focused (validity and educational impact), relevant and contextually appropriate (practicality and acceptability), standardised marking sheets/rubrics should be used (validity and reliability), and they did not tolerate poor quality assessment practice, such as the arbitrary or unjustified adjusting of marks.

"I can honestly say that I've never set a paper to keep the students' marks in a ... you know, a good range, one that would be acceptable within the rest of the faculty" (SA2).

(Standard setting, sophisticated assessment knowledge)

"I set my papers in such a way that I am convinced that it is a *fair* paper from the start … I would never just toss those questions for that specific paper … *I also don't believe in adjusting marks*, I think it is *a total abomination*… So this thing of, "Everyone has gotten 50% so let's push the marks up to get an average of 64%" is absolute nonsense" (SA1) – *emphasis* added.

(Standard setting, sophisticated assessment knowledge)

"I think it's very difficult to draw conclusions from an isolated situation. I think *you need to look at a student over a broader perspective* … Getting *a broader overview of the student's ability* and knowledge and skills rather than it being a single day for an assessment – a one off assessment" (SA11) – *emphasis* added.

(Validity, reliability, sophisticated assessment knowledge)

"We have a rubric we use … in the study guide there is a rubric that they see, which we use to evaluate them on, evaluate different aspects of the questions" (SA2).

(Standardised marking sheets, sophisticated assessment knowledge)

"We have an assessment sheet that is standardized for all the presentations, which we give to the students beforehand, so when they're planning their projects they know what marks will be allocated to the different aspects and they are free to come and have a look at their assessment sheet" (SA9).

(Standardised marking sheets, sophisticated assessment knowledge)

The [assessment] paper has to be *representative of all of the work*" (SA1) – *emphasis* added.

(Validity, sophisticated assessment knowledge)

"[The assessment is] influenced by the *patient population*. So *our patients have dictated,* and with the *disease profiles that we see* in them, *they dictated what we ought to know*" (SA8) – *emphasis* added.

(Relevance/educational impact, sophisticated assessment knowledge)

An understanding of blueprinting was a part of the technical knowledge dimension of the Competency and Conceptually-focused/Application-directed conception of assessment, for it reflected their more advanced technical knowledge. These participants believed in curriculum coherence and alignment between all teaching, learning and assessment activities, as well as a more continuous form of assessment, in order for assessment to be valid.

"The first thing is that you'll have to define what the characteristics of ... what the level of knowledge is that one would expect of a first-year student, on that level, what is the level of knowledge of a second-year student, and what level of knowledge does a third-year student have, so one can determine a profile for each and then … you can set up your questions around that profile" (SA4).

(Milestones/blueprinting, sophisticated assessment knowledge)

"A continuous assessment of that person, and it includes everything: knowledge, ability, attitude ... the whole package, that you see [a student] developing before your eyes" (SA2).

(Continuous assessment, formative, student-centred sophisticated assessment knowledge)

"If we set up teaching modules, we have specific outcomes that we hope to achieve, in terms of knowledge, skills, and attitudes that we hope students will gain from and maybe in some ways, it's *a method of measuring if we've met our outcomes* … I think it's a *system* that would cover core knowledge and would be *consistent with the outcomes* of that module that people would have *planned on right at the outset* and an assessment that is fair in terms of the questions that are asked and something that covers the specific knowledge and skills and attitudes that you would have wanted to have it *achieved in that particular module*" (SA9) – *emphasis* added.

(Blueprinting, Miller's pyramid, validity, sophisticated assessment knowledge)

To summarise the **Competency and Conceptually-focused/Application-directed** conception of assessment, participants conceptually understood that assessment served many purposes, but, primarily, it was a tool to support the development and conceptual change of students, with the goal being to produce competent, safe and socially accountable doctors. Therefore, the dimension of purpose in this category extended from academic and summative (Content-focused/Reproduction-directed) to a focus on student learning, social and moral functions. Furthermore, assessment was seen to impact on students learning in a positive manner, driving deeper approaches to learning as lecturers and students adopted a long-term approach in order to develop competency and prepare students for future clinical practice (in contrast to the short-term view of assessment as a mere academic or psychometric or practical task where students aim to pass or achieve a particular grade). These participants also possessed a more sophisticated technical knowledge of assessment, displayed in their expression of concepts such as validity, reliability, educational impact, practicality and acceptability, for example, through mentioning elements of blueprinting and continuous assessment.

5.2.5. Pilot Study: Discussion of the Phenomenographic findings

Twelve final year/s medical course convenors at a South African HEI were interviewed and three conceptions of assessment were identified through Phenomenographic analysis: “Undirected”, “Content-focused/Reproduction-directed” and “Competency and Conceptually-focused/Application-directed”. Participants’ conceptions of assessment were found to be comprised of three dimensions: the purpose of assessment, the learning effect of assessment and their technical knowledge of assessment. Participants perceived varied purposes of assessment, ranging from largely pointless, to a tool for measuring factual reproduction, to a tool for supporting student learning and development, viewed the effect of assessment on learning as both negative and positive drivers of student learning, and either had a poor, developing, or more advanced technical knowledge of assessment.

For the pilot study, archetypes of a participant for each conception are illustrated with quotations (see **Table 5.2.**). A participant with an "Undirected" conception would be "**SA6**" who had a negative view of assessment and viewed it as a way to punish the students. This participant possessed a poor technical knowledge of assessment, which was observed in the adjustment of student marks without qualm or question. A participant with a "Content-focused/Reproduction-directed" conception of assessment was seen in "**SA3**" as this participant viewed assessment as a way to get students to engage with and reproduce the content taught by the teacher, which was then forgotten after the assessment, which displayed a short-term and superficial learning effect. This participant also held an under-developed technical knowledge of assessment, seen in their current unawareness of concepts such as validity and reliability. The participant with a "Competency and Conceptually-focused/Application-directed" conception was exemplified in "**SA4**" who believed that the purpose of assessment was to prepare students for safe future clinical practice, and, therefore, designed assessments with questions that tested student insight and understanding. This lecturer also mentioned continuous assessment, feedback and hinted at concepts of blueprinting, milestones and programmatic assessment, which revealed an implicit and advanced awareness of technical assessment knowledge.

Table 5.2.: Pilot study: Archetypes of lecturers' conceptions of assessment in a single South African context.

Undirected (SA6)	Content-focused/ Reproduction-directed (SA3)	Competency and Conceptually-focused/ Application-directed (SA4)
"A necessary evil" (SA6) *"How well students respond to the threat of a test" (SA6)* *"[Assessment] is purely for a short-term goal and has no real meaning for the student in terms of the following year" (SA6)* *"Basically, again it is the … the stick and the carrot. Or rather the stick and not a lot of carrot" (SA6)* *"I just compensated by arbitrarily … by hiking up their marks … The assumption that the average mark would be somewhere around about sixty-five, and of course that's a complete assumption, [but] if you see marks that are substantially less than that then you can adjust*	*"[Testing the] minimum level of factual knowledge" (SA3)* *"The superficial, factual knowledge will be forgotten after the exam … Realistically a big part of the material that is studied is forgotten … I think that if we're honest with ourselves the majority of students will focus on memorisation of the facts" (SA3)* *"It's [assessment] a mechanism that forces the students, if you can use that word, to … study … It is a sort of external drive … the necessity to get them to study" (SA3)* *"For me it I s… the most important thing for me, if you can remember that [Professor X] said" (SA3)* *"I like to think that by asking different students' different things, I can gauge an overall impression of their place on the marks scale. However, I*	*"Assessment, in the end, has to be a reflection of … how the student has grown in terms of that subject field; how he has developed" (SA4)* *"Insight … deeper level … [Assessment] has to actually be in line to help students improve their skills or their abilities to become better doctors" (SA4)* *"A clinical problem scenario … a patient gets sketched for them every time. They are brought to a virtual space which reflects how their practice might look someday … important skills that a doctor has to have" (SA4)* *"Forms of continuous assessment … What we attempted was to try and push students in the direction … of self-, of active study" (SA4)* *"To offer feedback to a student about his knowledge and to indicate where the defects lie … Actually, the ideal is that the students can be reflective about their assessment" (SA4)* *"Higher order questions … insight … relevance … I think the more clinical the questions are, the more scenario-based ... case studies … It forces them [students] to integrate the work … There has to be a few questions that differentiate between students who have done in-depth study and the others … Students who show insight are those who end up faring the best" (SA4)*

those or not as you see fit" (SA6)	*know it's certainly not something one can prove scientifically … It's a very subjective thing … I have to say it's a gut feeling … There is no objective or quantitative measuring instrument I use" (SA3)*	*"The first thing is that you'll have to define what the characteristics of ... what the level of knowledge is that one would expect of a first-year student, on that level, what is the level of knowledge of a second-year student, and what level of knowledge does a third-year student have, so one can determine a profile for each and then… you can set up your questions around that profile" (SA4)* *"A radical way of doing this would be to assess them when they don't even know they're being assessed" (SA4)*

As the focus of this study is lecturer assessment behaviour, it was of interest to determine whether or not there was a potential causative relationship between conceptions and practice.

5.3. Pilot Study: Lecturer Assessment Practice in a South African context

Three conceptions of assessment were identified during the pilot study, ranging from Undirected to Content-focused/Reproduction-directed to Competency and Conceptually-focused/Application-directed.

Importantly, caution must be practiced when making potential claims linking conceptions to practice, for claiming an association or causal relationship between said conceptions and practice is beyond the scope of this study. However, I am cautiously suggesting that I have observed a link between conceptions and practice in this study, referring to this as the relationship between espoused (held) and enacted (practiced) conceptions. It appeared as though participants with Undirected conceptions enacted an "**Undirected assessment practice**", Content-focused/Reproduction-directed participants enacted a **"Content-focused/Reproduction-directed assessment practice"**, and Competency and Conceptually-focused/Application-directed participants enacted a **"Competency and Conceptually-focused/Application-directed assessment practice"** (see **Table 5.3.**). While these "categories" of enacted conceptions share the same category titles as the pilot study Outcome Space, this table is not an Outcome Space, and these are not dimensions of said conceptions. Rather, they merely provide practical manifestations of the conceptions and their variations, and give additional support to the categorisation of these conceptions, and organisation of the Outcome Space. These practical manifestations were also used to identify conceptions, whereby the researcher interpreted and untangled underlying conceptions from descriptions of assessment practice.

Briefly, assessment practice here refers to the assessment formats, tools and methods used by lecturers, including the types of questions asked, and the general assessment design process, reasoning and decision-making.

Table 5.3.: Pilot study: Lecturer Assessment Practice in a single South African context.

Undirected	Content-focused/Reproduction-directed	Competency and Conceptually-focused/Application-directed
• Undirected: little thought given to assessment practice • Negative: punitive	• Focus: content reproduction = academically-focused • More simplistic: traditional	• Focus: knowledge application, clinical competency and student development = socially-focused • More sophisticated: insight and innovation (clinical-scenarios, multiple formats and tools, feedback provided)

5.3.1. Undirected assessment practice

The **Undirected** conception of assessment consisted of a largely purposeless and negative view of assessment. These participants used assessment punitively, for example, as a form of punishment and a tool to inflict suffering, in the form of inducing stress, anxiety and fear in students. Due to the poorly developed assessment illiteracy of these participants, their assessment practice was often of a subjective nature, displayed in a lack of standard setting and the adjustment of marks, in stark contrast to the technical principles of assessment practice. In general, assessment was perceived to be a nuisance and burden, a practical task that simply had to be done, and, as such, convenience and ease of practice were prioritised.

> "An examination or an evaluation, it doesn't matter what it is [called], [it] is never a one hundred percent pleasurable experience. It can't be because you know it is … not an examination then, it isn't an evaluation then" (SA1).
>
> (Negative perception, punitive)

> "Sometimes they [students] deserve a little bit of punishment" (SA5).
>
> (Negative perception, punitive)

"The ease of which you can mark your paper" (SA11).

(Negative perception, burden)

"It's just sort of *gut feel*… You want something that's *easy to mark*, quite frankly, that's probably the most [important] ... *How easy is it to mark a question*" (SA12) – *emphasis* added.

(Negative perception, burden, little effort, subjective, no or incorrect assessment knowledge)

"I just compensated by arbitrarily… by hiking up their marks… The assumption that the average mark would be somewhere around about sixty-five, and of course that's a complete assumption, [but] if you see marks that are substantially less than that then you can adjust those or not as you see fit" (SA6).

(Subjective, uncritical, no or incorrect assessment knowledge)

5.3.2. Content-focused/Reproduction-directed assessment practice

The **Content-focused/Reproduction-directed** conception of assessment consisted of conceptually understanding assessment both concretely (as a short-term, practical task) and abstractly (as a potentially formative tool). These participants believed that assessment served the purpose of measuring student knowledge (a summative function) and as a tool to motivate student learning (a formative function), but, mostly stressed a content-focused and reproduction-directed assessment practice. This conception was more teacher-centred (as a content expert tasked with ensuring students know important information) than student-centred (in terms of prioritising their learning). The emphasis of this assessment practice was not the conceptual development of a student but rather more about the measurement of information in the assignment of a score that reflected how much a student was able to recall.

While studies have shown MCQs to be valid in testing student insight and understanding, in this instance, due to the simplistic technical assessment knowledge of these participants, MCQs and

short-answer questions were used to measure the reproduction of factual information over encouraging deep thinking and application of knowledge.

> "We use your *purely knowledge based type of questions*, say 'name' or 'list', or, you know, those types of pure knowledge-based questions" (SA2) – *emphasis* added.
> (Content reproduction, simplistic practice, poor assessment knowledge)

> "My impression is that there is a trend … [that] the questions can get a bit too … *factually orientated*, rather than being orientated towards comprehension" (SA3) – *emphasis* added.
> (Content reproduction, academic-focus)

> "We don't have a way of asking how important ... how the thought process ... that is not really tested … [The] students will get to the right answer because … they would've memorised the lecture … It's just *facts* that they [lecturers] were looking for" (SA4).
> (Content reproduction, superficial learning)

> "*Very little insight questions* are actually possible … You can't really test insight … A lot of it has to be *knowledge*" (SA5) – *emphasis* added.
> (Content reproduction, simplistic practice, poor assessment knowledge)

5.3.3. Competency and Conceptually-focused/Application-directed assessment practice

In line with the **Competency and Conceptually-focused/Application-directed** conception of assessment, these participants practiced assessment summatively, formatively, socially and morally. The emphasis of their assessment practice was to guide students in their progress and development in preparation for their future clinical work to ensure safe patient care. This was seen in their assessment practices which focused on clinical scenarios that required the application of knowledge.

"What I basically did was look at *different methods* of setting exam papers… as well as getting *integrated thinking* involved – *application of knowledge* … We used *clinical scenarios*, as well as *simulated patients*, as well as the *skills-lab models*" (SA11) – *emphasis* added.

(Knowledge application, clinical and patient-focus, competency, varied practice, sophisticated assessment knowledge)

"I think the more *clinical* the questions are, the more *scenario-based*, and we used quite a few *scenario*-based ... *case studies*… to *integrate* the work" (SA4) – *emphasis* added.

(Clinical and patient-focus, competency, long-term)

"They've got *understanding* about some of the concepts. We give them an abstract and say what does the following *mean* and what do they *think* … how would they *interpret* that to a *patient* who says *why* are you giving me this drug and this study shows something completely different. Or give them a newspaper article and say what actually was study was done here and how would they measure what was actually going on, what would you say to your *patient*. It is very much that kind of … it's very much *apply* [your knowledge]" (SA12) – *emphasis* added.

(Knowledge application, patient and social-focus)

Their more advanced technical knowledge of assessment was displayed through their use of varied assessment methods, specifically chosen to fit the function. They also explicitly mentioned their use of blueprinting, continuous assessment, standardised rubrics and feedback, which reflected a developed awareness of quality assessment practice.

"I think it's a system that would cover core knowledge and would be *consistent with the outcomes of that module that people would have planned on right at the outset* and an assessment that is fair in terms of the questions that are asked and something that covers

the specific knowledge and skills and attitudes that you would have *wanted to have it achieved in that particular module*" (SA9) – *emphasis* added.

(Blueprinting, validity, competency, student-centred, formative, sophisticated assessment knowledge)

"I would say that *frequent* assessment is important, because the more you force the student to engage with the information, the better, rather than him simply having these periods where he crams the information now, now, and now to get it together … to learn, to keep up by means of *continuous assessment* … The *more often* these class tests occur the more the students stand to gain from it … I feel that the *increase in knowledge* amongst students and the *application* thereof is something that should be *developed step by step*, it's *a continuous process*, but you take it step by step by step" (SA2).

(Continuous assessment, reliability, formative, sophisticated assessment knowledge)

"A radical way of doing this would be *to assess them when they don't even know they're being assessed*, but then they'll be assessed at that point based on the level of knowledge they have to have that time … And we say, "You're a student on this and this level who has had this and this exposure, and at this stage this is the type of product I want to have." And … we measure him based on … what he has to be at that stage, not what he's been doing the past four weeks, but based on the past four years … The first thing is that you'll have to define what the characteristics of ... what the level of knowledge is that one would expect of a first year student … what is the level of knowledge of a second year student, and what level of knowledge does a third year student have, so one can determine a profile for each and … then you can set up your questions around that profile" (SA4).

(Milestone/blueprinting, student-centred, formative, sophisticated assessment knowledge)

"I think one has to move towards *continuous evaluation* slowly but surely" (SA1) – *emphasis* added.

(Continuous assessment, reliability, sophisticated assessment knowledge)

"[We practiced] forms of *continuous assessment* … What we attempted was to try and push students in the direction … of self-, of active study" (SA4).

(Continuous assessment, formative, student-centred)

"To offer *feedback* to a student about his knowledge and to indicate where the defects lie … Actually, the ideal is that the students can be reflective about their assessment" (SA4) – *emphasis* added.

(Formative, student-centred)

"The assessment would be *a clue*, [a] student is struggling, therefore someone needs to make sure that they get through … the student would go back and then redo the assessment, and the assessment then would then be a key to increasing skills and knowledge. So that ... you wouldn't get through, it wouldn't give you a great mark, but would clarify understanding … I think if you want to change the assessment so that it becomes more *ongoing in a dialogue kind of way*" (SA12) – *emphasis* added.

(Formative, student-centred)

To summarise, I tentatively suggest that there appeared to be a possible link between participants' conceptions of assessment and their assessment practice, or the espoused and enacted conceptions, which I am calling **Undirected assessment practice**, **Content-focused/Reproduction-directed assessment practice**, and **Competency and Conceptually-focused/Application-directed assessment practice**. Determining the possible strength of association between said conceptions and practices remain beyond the scope of this study.

The data collected and analysed in the pilot study represented a single medical HEI in South Africa, with a particular sociocultural heritage and disposition (traditionally white and Afrikaans in language and culture). In order to further explore and validate the proposed conceptions of assessment, data was collected and analysed from additional Southern sites, beginning with a second South African HEI (traditionally white and English in language and culture, but currently more representative of the South African demographic), followed by data collection and analysis

at a single HEI in Mexico (with a Spanish colonial heritage). In the next chapter, findings from the second and third phases of this study are reported.

Chapter 6: Confirmatory study: Lecturers' Conceptions of Assessment at additional Southern contexts

A number of conceptions of assessment held by final year/s medical course convenors at one South African site were identified and described in the first phase and pilot study of this research project. In order to confirm said conceptions, further data collection and analysis took place at two additional diverse Southern contexts (a further **17 interviews** totaling **105 000 transcribed words**). Six lecturers were interviewed at a second HEI ("SA13-18") in South Africa (traditionally English in language and culture, in contrast to the previously Afrikaans HEI, yet both of a Dutch and British colonial history, but independent and diverse today), and thirteen lecturers were interviewed at an HEI in Mexico ("MX1-13", of a Spanish colonial history and independent today).

It must be noted that at the time of developing the second Outcome Space, data was first analysed inductively, considering the second South African dataset alone, before deductively consulting the findings from the pilot study. This was followed by data collection and inductive analysis of the Mexican dataset alone, before deductively looking at the previous two South African datasets, which then informed the third Outcome Space. The first and second Outcome Spaces simply reflect the ongoing processes of data collection, analysis, understanding and theorisation. The third Outcome Space is the final Outcome Space, for it is representative of *all* the data collected and reflects the deepest level of interpretation and abstraction in this study.

6.1. Second phase: Lecturers' Conceptions of Assessment at a second South African context

The confirmatory study consisted of the second and third phases of data collection and analysis, referring to the second South African and Mexican datasets respectively. During the second phase of this study the second Outcome Space was developed, which consisted of three conceptions of assessment and six dimensions. The conceptions were titled "**Detached practitioner**", "**Emerging equilibrium**" and "**Engaged educator**" and the dimensions were labelled "purpose", "assessment literacy", "temporal perspective", "identity" and "role" (see **Table 6.1.** for an outline and **Appendix 6.1.** for a table with illustrative quotes).

6.1.1. Second phase: The Dimensions of Lecturers' Conceptions of Assessment

As previously described, dimensions refer to important aspects of a phenomenon, and thus, components of lecturers' conceptions of assessment. In hindsight, it was noted that the initial dimensions identified in the pilot study were prescriptively based on the content of the interview questions. Reflexive, careful and critical openness was then practiced during the coding steps of Phenomenographic analysis during the confirmatory study. Both the second South African and Mexican datasets were first analysed inductively, in isolation, before, deductively considering the other datasets available.

In the second Outcome Space, three conceptions and five dimensions were identified. While an increase in the number of dimensions occurred, there was overlap and coherence between the first and second Outcome Spaces. For example, both Outcome Spaces shared the dimensions of "purpose" and "knowledge". "Technical knowledge", in the pilot study, was changed to "assessment literacy" in the second Outcome Space, for it better reflected the broader awareness and understanding participants had towards assessment, as opposed to the technical aspects alone. The dimension of "learning effect" was incorporated into the "purpose" of assessment. Purpose was also linked to the new dimension of "temporal perspective", for how

one viewed the purpose of assessment implied their short- or long-term perspective of assessment. Unique dimensions that emerged during the second phase of data collection and analysis were “identity” and its related “role”.

The code “assessment literacy” ranged from incorrect and poorly developed ideas of assessment, to implicit craft knowledge, that was based on clinical experience and expertise developed over time, to a more advanced and scholarly knowledge on the other end of the spectrum, evidenced by explicit descriptions of key assessment principles and techniques. This dimension was displayed either explicit articulations or descriptions of their assessment practice, which I interpreted as illustrations of their assessment literacy. For instance, if a participant mentioned a preference for the use or various assessment tools and a continuous assessment, I could interpret that, based on additional comments and the transcript as a whole, this participant possessed a more sophisticated understanding of high-quality assessment practice and the principles that underpin it, such as validity and reliability.

An interesting dimension that emerged during the second phase of this study was “identity”. This came about through noticing that the participants often directly referred to themselves as “clinicians”, indicating the importance of this point. While identity did not emerge as a dimension in the pilot study, it was also not asked of the participants, whereas, as it was strongly emphasised in the first interview of the second phase, it was then asked in subsequent interviews to determine whether or not it formed part of these conceptions. Identity in this study, as has previously been described (see **Section 3.3.1.2.**), refers to how a lecturer sees themselves, as a clinician (“practitioner”) or “educator” or negotiating a balanced mixture of the two (discovering duality and working towards establishing an “equilibrium”).

While “identity” was not present in the pilot study Outcome Space, I could argue that aspects of it were perhaps implicitly present. For instance, linked to “identity” is the dimension of “role”. Based on how lecturers see themselves influences how they conceptually understand their role in assessment. For instance, a participant who identified strongly as a clinician, valued and

prioritised their clinical work, and viewed assessment as an onerous task. They would then adopt a more passive role, such as merely implementing and managing assessment, without much thought or effort. As these participants did not value assessment, took a negative stance towards it, and gave it *no direction* from themselves, this was reminiscent of the "Undirected" conception of assessment.

The "Detached practitioner" conception of assessment privileged disciplinary knowledge, and the role the practitioner played in imparting this knowledge, which was similar to the "Content-focused/Reproduction-directed" conception of assessment. A participant who identified as a "Engaged educator", conceptually understood their role to be that of a facilitator of student growth and development, and was active in the design and initiation of their assessment practice, for they saw it as important in shaping student learning to lead to the desired outcomes of real competence for the sake of their future patients. This was displayed in the "Competency- and Conceptually-focused/Application-directed" conception of assessment. I believe that the dimensions and categories from the first Outcome Space were adequately incorporated in the second, and, later, the third and final, Outcome Spaces.

As identity formed a foundational theme in the second phase of the study, so it was used to label the categories of description, as lecturers often referred back to their core identity as a strong justification for their assessment decisions and practices. At this stage of the qualitative research process, identity also seemed to adequately account for, and link to, all the dimensions within each conception. However, as analysis deepened and enriched, so the titles of the conceptions in the third and final Outcome Space were changed.

In line with the relational aspect of Phenomenographic Outcome Spaces, the relationships between dimensions has been described above. This also relates to the logical organisation of the Outcome Space, for dimensions were sorted together according to how they refer and relate to one another within individuals and across transcripts.

Table 6.1.: An outline of the Phenomenographic Outcome Space outline reflecting lecturers' conceptions of assessment in a second South African context.

	Detached practitioner	**Emerging equilibrium**	**Engaged educator**
Purpose	• Assessment drives learning (summative – administrative) • Clinical competency goal (teacher-centred, disciplinary knowledge-centred) • Concrete: practical task	• Assessment drives learning (both: formative and summative) • Both: concrete (administrative) and abstract (learning)	• Assessment drives learning (formative) • Student learning and development goal (student-centred) • Clinical competency goal (moral and social focus) • Abstract: learning and developmental focus
Temporal perspective	Short-term (administrative task)	Both: short- and long-term	Long-term (formative development and clinical competency)
Assessment literacy	• Incorrect or poorly developed • Implicit craft knowledge: content/disciplinary experience and expertise	Both: implicit (clinical expertise) and emerging awareness of HPE evidence-based practice	Advanced understanding and use HPE evidence-based
Identity	Practitioner (clinician in this context)	Both: clinician and educator	Educator
Role	• Role: administrative, manager, operator (pragmatic or mechanistic view: simply implements) • Content/disciplinary knowledge and expertise privileged	• Role: both a practitioner and education, developing a negotiated/balanced identity) • Establishing equilibrium between content/ disciplinary expertise with educational training (see both identities as important and necessary)	• Role: educator and assessor (initiates and in full control/ownership of assessment practice) • HPE evidence-based practice privileged

6.1.2. The Detached practitioner conception of assessment

The first major conception of assessment of the second phase Outcome Space was titled **"Detached practitioner"**, as central to this conception was the understanding that these individuals conceptually understood their primary identity and role to be practitioners (clinicians in this context) and not as educators, or more specifically, as assessors.

These individuals used their identity as clinicians as a justification to excuse their lack of assessment literacy, rather, their knowledge and expertise lay in their clinical disciplines. As clinicians, whose clinical workload was valued and prioritised, they lacked a desire to practice assessment. As such, they felt no sense of ownership towards their assessment practice, in other words, they were detached practitioners.

> "So, I just got involved [in assessment] because it *fell* [into my lap], I was *delegated* to be the course convenor" (SA13) – *emphasis* added.
> (Detached, no ownership)

> "We just *rotate through* the department whose turn it is to examine" (SA13) – *emphasis* added.
> (No ownership)

> "I'm a *clinician*, I am not [a teacher], I haven't got any experience in education and my *main interest* is *clinical work*" (SA14) – *emphasis* added.
> (Practitioner identity)

> "I am a *clinician first of all* … So, I am a *clinician*, okay? ... We are all *clinicians* we are all doing this as *extra*, you know. So, I have a *full clinical load* … I have got other responsibilities" (SA16) – *emphasis* added.
> (Practitioner identity)

> "The vast majority of convenors and examiners in the clinical years are *clinicians* … The vast majority of the teachers are *first and foremost clinicians* and their teaching commitments are *added on top*" (SA18) – *emphasis* added.
>
> (Practitioner identity)

Assessment was seen as a task, something concrete or practical, viewed negatively, as a burden, because it took time away from their main and more important work as clinicians.

> "Our *main job* is not doing assessment, it is running the ward and doing other things, and it [assessment] is sometimes seen as a *chore* we have to do, especially if it is every two weeks. So, it is … something you just have to *get over with*, so it is *not* always the most *effort* is put into it … not saying that I have given it more thought … [Assessment] is very much a *little extra thing* that we do, it is not our main focus" (SA13) – *emphasis* added.
>
> (Practitioner identity, concrete task, no ownership, disciplinary knowledge)

> "The burden of assessment" (SA14).
>
> (Concrete task, negative perception)

For these participants, the <u>purpose</u> of assessment was administrative. The <u>short-term</u> focus of their assessment practice was to obtain a score for each student by the end of their course. In line with this administrative view was the pressure to pass students, which led to a defensive-type of feedback being given to students, not to assist them in their development, but rather performed to avoid and protect themselves from further administration, such as having to organise for a student to repeat a block or preempt possible queries by students regarding their assessment scores. Their concerns were of a pragmatic nature, and an avoidance motivation goal (to prevent a negative effect or outcome, in this case additional administrative work). This was also observed in their passivity and lack of assessment ownership, striving for convenience and seeking to put in minimum effort, and a more negative perception of assessment.

"They have to know their stuff by the following Friday … It is just because we are *required* to assess them as part of their *mark*" (SA13) – *emphasis* added.

(Concrete task, administrative, short-term perspective)

"[Assessment must] motivate them [to study], because if they don't get *a pass* then they have to repeat the block … They [students] need to *pass that exam* and then their overall mark must also be *a pass*" (SA14) – *emphasis* added.

(Administrative, short-term perspective, pass focus)

"We do (give) feedback after the exam, what *marks* they actually got and any feedback, what things they have done well and things they haven't done well. We do it straight afterwards to *prevent*, sort of, down the line someone *querying a mark* … The consultants will ask them if they want to *see the marks* and discuss any problems that they have had, so, that they can *address any issues right there* than come back months later when no one can actually remember ... Because I am the convenor, I do get the occasional student sending me an email or coming to me, that the student is upset, that they should have gotten a higher *mark*, and then you are trying to sort that out months later is *a bit of a mission* because nobody can remember much about what happened. So, it is only this year that we decided to give students feedback immediately and then *deal with any issues* straight away" (SA13) – *emphasis* added.

(Administrative, short-term perspective, pass focus, negative perception)

"To be honest, another thing is if the student fails the exam, it is quite an *administrative nightmare* to then try and sort that out and to make sure that they come back and do re-exams. So, it is an *incentive* to not let them fail and if they are really bad you'll perhaps try and find a way to give them *at least 50%* so that you won't have to (deal with the administrative nightmare) … But having said that, if somebody is terrible, I don't think we would do that. So, it is just for the borderline cases you might, because of the *administration issue*" (SA13).

(Administrative, pass focus, negative perception)

The assessment practice of these lecturers, and their role in it, could be described as administrative, managerial, mechanistic or even operational, because they merely managed the physical assessment task "on the ground" with little thought or effort – it was simply required, routine practice. This related to a lack of felt responsibility and ownership towards assessment. This practical task-driven, passing students along, approach to assessment revealed a short-term, "today", perspective, which aimed to pass students through the clinical block, as opposed to a long-term, "tomorrow", perspective, which would focus on competency development for safe future practice and patient care.

> "As the *organiser* for the students at [Hospital X], I can't speak for the other hospitals… If you want to speak to the other two hospitals [clinical rotation sites] you will have to speak to them … I would assume [it] would be *someone else's responsibility*" (SA13) – *emphasis* added.
>
> (Operational, detached, no ownership)

> "I tend do to the sort of *managing* the students *on the ground*, a lot of these *nuts and bolts* of the programme" (SA16) – *emphasis* added.
>
> (Operational, administrative, concrete task)

These lecturers possessed a poor assessment literacy, reflected in their lack of awareness of foundational assessment principles, as observed in their passive, unquestioning and uncritical acceptance of current ("inherited") assessment tools, such as using a standardised marking sheet (rubric) that already existed when they took over as course convenor and assessment-lead.

> "So, no sort of scientific … way of weighting it … Initially it was out of fifty marks, now it is out of a hundred" (SA17).
>
> (Unaware, poorly developed assessment literacy)

> "I *inherited* [a rubric] from the *previous* convenor … I am *not sure* exactly where he got it from, I think he probably adapted another … So, we basically have been using it, but I

don't really know where it comes from … I *haven't actually thought about* doing my own one" (SA13) – *emphasis* added.

(Operational, uncritical, poorly developer assessment literacy, no ownership)

"We actually have a proforma [rubric] … I *don't know* who's come up with this specific one, but I think there has been multiple iterations over the years and this is the current version" (SA16).

(No ownership, uncritical)

If anything, these participants relied on their clinical expertise and experience, a kind of craft knowledge, for assessment-related decision-making, over using HPE-derived objective measuring tools, such as standardised marking sheets. These participants privileged their clinical disciplinary knowledge and expertise and used them as the foundation for their assessment-related decision-making, and allow those who assess during their course to continue this practice.

"Although we have this proforma, this form, often we find it easier to just not, sort of use it … We have done it enough to know what percentage to give, the rough mark to give the student and generally most of the examiners won't fill this in in detail, they will just give an overall mark … Although, we have others who use it perfectly … [But] sometimes just an average score will be given as well. And sometimes, it's done a few weeks after they have actually been there and you can't actually remember them that far, so, again, it is not perfectly done … It's mainly just from our own experiences … plus getting to know the more you do … Most people [examiners] will just give an overall mark, which I suppose is a feeling that they get, how good the student is compared to the other ones that they have examined before" (SA13).

(Privilege practitioner identity and disciplinary knowledge and experience)

"Some examiners still don't like tick sheets, so they'll like listen to the story the student is telling them and then give a global impression without ticking … Some examiners just feel that-, and I don't even know if I think they are wrong, because you are so experienced

with something that just sitting there like a robot and ticking, you know, just giving your impression on how you think the student managed the patient is maybe a better way of doing it, but I can't argue for or against either one because I think that is challenging, especially when you have someone who is so experienced and you ask them to tick box … [It] is very subjective" (SA14).

(Privilege practitioner identity and disciplinary knowledge and experience)

"Most of the examiners … are also just clinicians who are working in the various hospitals … So, they are experienced but they are not trained" (SA14).

(Implicit craft knowledge, disciplinary experience)

"Almost all of the examiners, certainly at the undergraduate level, they would have done it many times … But many of them are also examiners attached to the [postgraduate] college, so, they are experienced at examining at both the undergraduate and postgraduate levels … All of our examiners are generally experienced" (SA18).

(Privilege practitioner identity and disciplinary knowledge and experience)

"The goal of assessment is to determine who actually has [disciplinary] knowledge" (SA16).

(Privilege disciplinary knowledge, content-reproduction)

To summarise, the "**Detached practitioner**" conception of assessment was based on the core conceptual understanding of participants as practitioners (clinicians), first and foremost. Their disciplinary knowledge, expertise and clinical workload were prioritised over assessment activities. This meant that their assessment literacy was lacking. These participants too saw assessment as a short-term, administrative, practical, burdensome task, which resulted in a rather passive (detached) approach to assessment, that aimed to pass students through their course with as little thought and effort from them as possible.

6.1.3. The Emerging equilibrium conception of assessment

The second conception of assessment identified was **"Emerging equilibrium"**. In this category, lecturers have a dualistic identity of *both* practitioner/clinician *and* educator. In speaking with these lecturers, it appeared as though these dual identities generally existed in a harmonious balance, for, while there was some negotiation and compromise, there was not a severe competition or tension, rather, both were seen to be important and valuable for their assessment practice (there was an "emerging equilibrium"). In fact, they proposed that in order to teach and assess at an HEI, academics should be *both* an experienced clinician, daily interacting with patients, and be well-equipped for their educational responsibilities.

While participants with the "Detached practitioner" conception of assessment were passive in their approach to quality assessment principles, such as using standardised assessment tools given to them, primarily relying on their experience and expertise as clinicians for decision-making, participants with conceptions of "Emerging equilibrium" displayed a developing and implicit assessment literacy, perhaps due to time spent assessing and critically reflecting, becoming "self-taught" through their gained craft knowledge of quality assessment practice, or through having been exposed to (or seen the need for) HPE and assessment training. A level of activeness was observed in participants with "Emerging equilibrium" conceptions as was observed in their questions and grapples with technical issues, and their engagement in these matters through expressing their uncertainties instead of standing confidently behind their clinician identity and disciplinary knowledge.

> "If a student is pass/fail, that is really where we will *interrogate things quite closely* … We don't lower marks, we can't do that … [but] we may *interrogate* the oral and say, "Okay, you failed the student, are you sure? Why did you fail the student? Is that appropriate or accurate?"" (SA18) – *emphasis* added.
>
> (Questioning, developing awareness and assessment literacy)

"We've spent the last two years to *really try* and *fine tune our assessments*, to make it very *objective*, and I think we are getting there. Our results seem to be a lot more realistic. We did go through a period of time, last year, the year before, where the marks were very high… I guess that means the students are really bright or maybe we are marking too leniently … So, this year, definitely, we can see there is a change, we are not actually failing more students, but the average marks are a lot more realistic, and we are able to tease out the good students" (SA14) – *emphasis* added.

(Emerging, becoming more active, developing awareness and assessment literacy)

"One speaks about alignment and how important it is, but often one isn't taught how to check for alignment – how do you measure alignment? Is there a tool for saying, "Yes, this is a well-aligned course or programme or no it isn't"? ... So that kind of training and ongoing capacity development around assessment … I think more of that could be happening. You know, you think of new staff coming in and need to design an assessment … it's quite challenging setting exam papers or … designing assessment rubrics … When I was designing exam papers … it was mostly from my own prior reading and learning; and *then one just develops skills over time*" (SA17) – *emphasis* added.

(Aware, active, developed assessment craft knowledge)

"I can never remember the terms, formative and… [summative]? I can never remember and I always get confused about which is which" (SA18).

(Emerging, developing awareness)

"I have just kind of been *exposed to different things as I've moved along*. I'm involved, I'm on the assessment committee of the university, I've examined at all levels of the college" (SA18).

(Developed assessment craft knowledge)

In terms of the other Outcome Space dimensions, role and temporal perspective, the "Emerging equilibrium" category shares and possesses dimensions from both the "Detached practitioner" and "Engaged educator" conceptions, for participants who hold the "Emerging equilibrium" conception of assessment see it both as a short-term, concrete, administrative task and as a long-term, abstract, developmental tool for student learning – these participants are trying to establish equilibrium, because they believe both identities are important and necessary for a quality assessment practice.

In summary, the "**Emerging equilibrium**" conception of assessment was unique from the other conceptions in that these participants possessed both identities and viewpoints of a practitioner/clinician and an educator. Their challenge was to balance these dual identities, perspectives and roles (achieve "equilibrium"). Similarly, in terms of their assessment literacy, they saw the value and importance of both craft knowledge, based on clinical experience and expertise, as well as HPE training, seen in their awareness of basic concepts of quality assessment, which was also expressed through their active questioning and developing understanding of assessment concepts. However, at this stage, they do not yet hold advanced or explicit conceptions of said technical assessment knowledge.

6.1.4. The Engaged educator conception of assessment

The third major conception of assessment from the second phase of data collection and analysis was **"Engaged educator"**. The primary identity of these participants was that of an educator, someone who valued and prioritised their professional development and further education as an educator, and their students' development. This included seeing their assessor role educationally too, for assessment was an extension of their educator identity, for it is used as a formative tool to guide student learning. This reflected a long-term perspective in preparing students for future clinical practice and patient care, which also revealed a social and moral purpose.

"In their log books they have to do *formative* case presentations … and they get *feedback* from the consultants in terms on *how they can improve* and it *doesn't count for a mark* … It is purely for the *student's benefit*, there is *no mark* that goes along with it and the students are encouraged to use that as an *opportunity to learn*" (SA14) – *emphasis* added.
(Educational focus, assessment for student learning, long-term perspective, sophisticated assessment literacy)

"*Reflecting* on their learning objectives … their clinical activities, the *reflection* on their *professional and personal development* … To what extent they [students] felt the [course] helped to *promote agency* … you know, as *future interns*, how will they … function? Do they feel prepared, energized, inspired, an *agent of change*? … Once they [students] sat down to write it [reflective report] they realised how much they had *learned*, for some it is a revelation that they knew more than they thought they did" (SA17) – *emphasis* added.
(Assessment for student learning, moral and social purpose, long-term perspective)

"The *formative promotion of student learning*, where the *feedback helps to improve* … The informal *feedback* … They [students] [are] welcomed as part of the *health team*, being given the *responsibility* to do procedures that they ordinarily wouldn't be able to do, feeling like they are *functioning as a doctor* than just a student in this big hierarchy … That's as valuable for their *formation as … professionals to practice* as the formal feedback on the report" (SA17) – *emphasis* added.
(Assessment for student learning, moral and social purpose, long-term perspective)

"Who actually has *knowledge*, who has *good clinical skills*, who is *safe*, okay? So, we are trying to make sure that there is a *certain standard of knowledge or certain level of knowledge and clinical abilities*. So the assessment of that knowledge and their clinical skills is what we are trying to achieve with the end of block examination" (SA16) – *emphasis* added.
(Knowledge application and competency, social and moral purpose)

"If you have [formative] assessments during the course of the block that … don't contribute to the final mark, then that obviously is of great help … [and that] removes the anxiety around the sort of high stakes exam that you either pass or fail … That's something I'm working towards, but, at the moment, I am still building my question bank to use for the formative assessment … Once I have sufficient clinical material there I can start expanding and building questions that can be used during the course. So, that is something that is in the pipeline" (SA18).

(Assessment for student learning)

"I mean it's natural that one fears an examination – it's unfortunate … And it's of some concern to me to see the anxiety that candidates come with, and it's across the board … Well, you know, it's a high stakes event … [and] failing an examination has massive implications. So, people fear that and so they learn, they study, so that they pass their exam. If you didn't have an exam, then people might not work as hard, so that's one concept of what assessment driving learning. But … it obviously has other sort of interpretations as well … [At one HEI] there is no examination during the course and students still learn, so, it's not like people will not learn if there isn't an assessment. But, there is certainly a group of people who are probably going to spend more time with their books and learning because there is an exam … The other thing is, when one understands what is coming in an assessment, then that also drives how you learn… So, because students know there is a clinical exam with case vignettes and are expected to make a diagnosis and an investigation plan and a treatment plan, then they start to think like that … and they will apply themselves in that way when they are exposed to that in their clinical environment. And I think that is probably more important when you talk about assessment driving learning … To see if they … are safe or not" (SA18).

(Assessment for student learning, social and moral purpose, knowledge application and competency, long-term perspective)

As educators, these participants possessed a more advanced assessment literacy, displayed through their explicit mention and practice of alignment and blueprinting. With their assessment literacy these participants were in greater control and ownership of their assessment practice, as observed in their active involvement from its conception to its implementation. For example, they selected different assessment tools to serve different assessment functions in order to meet specific outcomes.

> "The in-course assessment is by far the better reflection of what the students are worth because it is a longitudinal assessment over four weeks, [whereas] we are seeing them as a snapshot at the [final] exam. So you could pass an exam and not be safe" (SA16).
> (Awareness, critical thinking)

> "[Assessment] is really important, certainly from the student's point of view, for them assessment drives the learning, as we always hear, and so that is why it is so important that the assessment is properly aligned with the stated learning objectives and the activities" (SA17).
> (Sophisticated assessment literacy)

> "I blueprint everything. So, we have a curriculum, a core curriculum, that's our starting point … [That] outlines the learning outcomes and our students also have what we call a skills logbook. So, whatever they are assessed on is directly matched to the logbook which is based on the learning outcomes which can be traced back to the core curriculum … That's basically how we look at our assessments ... This whole thing about validity and reliability and all those things … [I] have the strong belief that you cannot just assess someone blindly … your starting point must be the core curriculum, they need to have this roadmap, "That is where you are going"; you need to have clear outlines of what you expect and you have to assess what they are exposed to. Whether they are exposed in the classroom or out in the workplace, so you have to be able to match, you can't assess them on things they haven't done" (SA15).
> (Sophisticated assessment literacy)

"I designed it [a rubric] … I do most of it [designing assessment questions] myself … Since I started convening the course, that is something I am always have been looking to more towards; there are quite a lot of changes I made initially … It's kind of a work in progress, to a large degree, I think we are always trying to work towards improving things" (SA18).
(Sophisticated assessment literacy)

"[Different assessment tools] each interrogate different things … and they all have their strengths and their weaknesses. Obviously, we want to ensure that they've covered the theory, so, and they've read the text and what we want them to know; so that's the point of assessment one. The point of assessment two is … to test their application ability; can they apply the knowledge to clinical scenarios … Assessment three, the oral and portfolio, make sure that we interrogate something that they've done during the block … and that they've read around the cases… The fourth exam interrogates actual practical clinical … skills … So, each of the formats does sort of assess different things and different aspects … There is no point having multiple assessment points but they are all the same; you need to decide what it is you are testing and focus your assessment on that … So, in other words, are you testing what you want to test?" (SA18).
(Sophisticated assessment literacy)

"We use paired examiners… and the idea would be that the examiners are intended to come to independent marks and then compare and come with a final mark … They [examiners] follow the rubric closely … They are expected to detail their interview [oral assessment]. We have an examiners meeting at the conclusion of the exams and at least one member of the pair of the examiners is expected to be at the meeting" (SA18).
(Implicit HPE principles)

In summary, the "**Engaged educator**" conception of assessment had a strong focus on using assessment as a formative tool to guide student learning and development, with the long-term goal being to ensure clinical competency. In order to do so, these participants possessed a higher-level of assessment literacy, displayed through their explicit use and practice of sound assessment criteria and techniques, such as using blueprinting to design valid and reliable assessments. These individuals saw themselves as active and responsible assessors.

6.1.5. Second phase: Summary of the Phenomenographic findings

In the second phase and confirmatory study of this research project, a second Outcome Space depicting lecturers' conceptions of assessment in the South was developed, consisting of three conceptions and six dimensions. Identity emerged as an important theme or category descriptor. Archetypes for the categories of "Detached practitioner", "Emerging equilibrium" and "Engaged educator" would be **SA13**, **SA18** (who displayed both/a mixture of "Detached practitioner" and "Engaged educator" conceptions) and **SA15**, respectively (see **Table 6.2.**). The extremes of the Outcome Space range were clear, "Detached practitioner" and "Engaged educator", however, the middle category of "Emerging equilibrium" was less clear, for it overlapped with dimensions from both categories on either sides and expressed elements of the "Detached practitioner" conception and aspects of the "Engaged educator", reflecting its developmental nature in working to establish a balance between the two.

Table 6.2.: Archetype of lecturers' conceptions of assessment in a second South African context.

Detached practitioner (SA13)	**Emerging equilibrium (SA18)**	**Engaged educator (SA15)**
"So, I just got involved [in assessment] because it fell [into my lap], I was delegated to be the course convenor" (SA13) *"I haven't actually thought about doing my own one [rubric]" (SA13)* *"As the organiser for the students at [Hospital X]" (SA13)* *"It [assessment] is just because we are required to assess them as part of their mark … It is what has always been done … It's a chore we have to do … something you just have to get over with, so it is not always the most effort is put into it … Our main job is not doing assessment, it is running the ward and doing other things" (SA13)* *"To be honest, another thing is if the student fails the exam, it is quite an administrative nightmare to*	*"The vast majority of convenors and examiners in the clinical years are clinicians … The vast majority of the teachers are first and foremost clinicians and their teaching commitments are added on top" (SA18)* *"I can never remember the terms, formative and…[summative]? I can never remember and I always get confused about which is which" (SA18)* *"I have just kind of been exposed to different things as I've moved along. I'm involved, I'm on the assessment committee of the university, I've examined at all levels of the college" (SA18)* *"We will interrogate things quite closely… We don't lower marks, we can't do that" (SA18)* *"If you have [formative] assessments during the course of the block that… don't contribute to the final mark, then that obviously is of great help … [and that] removes the anxiety around the sort of high stakes exam that you either pass or fail … I am still building my question bank to use for the formative assessment" (SA18)* *"It's a high stakes event… failing an examination has massive implications. So, people fear that and so they learn, they study, so that they pass their exam. If you didn't have*	*"I blueprint everything; so we have a curriculum, a core curriculum, that's our starting point … [That] outlines the learning outcomes and our students also have what we call a skills logbook. So whatever they are assessed on is directly matched to the logbook which is based on the learning outcomes which can be traced back to the core curriculum … That's basically how we look at our assessments ... This whole thing about validity and reliability and all those things … [I] have the strong belief that you cannot just assess someone blindly … your starting point must be the core curriculum; they need to have this roadmap, "That is where you are going"; you need to have clear outlines of what you expect and you have to assess what they are exposed to. Whether they are exposed in the classroom or out in the workplace, so you have to be able to match, you*

then try and sort that out and to make sure that they come back and do re-exams. So, it is an incentive to not let them fail and if they are really bad you'll perhaps try and find a way to give them at least 50% so that you won't have to [deal with the administrative nightmare]" (SA13)	*an exam, then people might not work as hard, so that's one concept of what assessment driving learning … There is certainly a group of people who are probably going to spend more time with their books and learning because there is an exam … The other thing is, when one understands what is coming in an assessment, then that also drives how you learn … So, because students know there is a clinical exam with case vignettes and are expected to make a diagnosis and an investigation plan and a treatment plan, then they start to think like that … and they will apply themselves in that way when they are exposed to that in their clinical environment. And I think that is probably more important when you talk about assessment driving learning" (SA18)* *"Each of the formats does sort of assess different things and different aspects … There is no point having multiple assessment points but they are all the same; you need to decide what it is you are testing and focus your assessment on that … So, in other words, are you testing what you want to test?"" (SA18)*	*can't assess them on things they haven't done" (SA15)* *"I did a two-year [HPE] fellowship in SAFRI [Sub-Saharan Africa-FAIMER Fellowship]" (SA15)* *"There is this whole thing about formative versus summative story and I am a firm believer that, especially in the clinical years, the best place to assess them would probably be in the workplace" (SA15)* *"We do review our programme every year, you know, we look at student feedback, which I think it important, and we try to improve – we try to improve. And, of course, I try to upskill myself in terms of, you know, educational things" (SA15)*

As in the pilot study, so a link between espoused and enacted conceptions was observed, lending support to the proposal that conceptions of assessment impact on assessment behaviour.

6.2. Second phase: Lecturer Assessment Practice in a second South African context

As with the pilot study, the same category titles of "**Detached practitioner assessment practice**", "**Emerging equilibrium assessment practice**" and "**Engaged educator assessment practice**", relating to the "Detached practitioner", "Emerging equilibrium" and "Engaged educator" conceptions of assessment, respectively, were used to describe the related assessment practice (see **Table 6.3.**). Clear examples include those with a "Detached practitioner" conception of assessment relied heavily on their clinical expertise and experience when grading students, while, those who held an "Engaged educators" conception, explicitly performed more sophisticated assessment practices, such as blueprinting and formative assessment.

The extremes of the Outcome Space spectrum, "Detached practitioner" and "Engaged educator", are clearer than the middle category of "Emerging equilibrium", which possessed elements of both conceptions and practice on either side. The "Detached practitioner" and "Engaged educator" practices will be described below, and the "Emerging equilibrium" practice only very briefly, as it was a combination of the two.

Table 6.3.: Lecturer Assessment Practice in a second South African context.

	Detached practitioner	Emerging equilibrium	Engaged educator
Goal	• Students to pass clinical block/ rotation (administrative and summative)	• Both/ mixture	• Student learning and competency development (formative)
Priority and Ownership	• Clinical work prioritised • Little ownership of assessment practice (merely implements rules and follows what others have done)	• Balance between clinical and educational work	• Explicitly engaged with educational work • Active ownership of assessment practice
Practice	• Mostly summative (administrative) • Traditional or status quo (do what has always been done) • Rules-based (follows policies, guidelines of HEI or department) • Requires help from others	• Both/mixture: uses clinical expertise and basic HPE • Mixture: sometimes initiates changes, sometimes requires help from others	• Formative focus • Uses multiple tools to achieve multiple purposes • HPE evidence-based practice • Initiates changes themselves

6.2.1. The Detached practitioner assessment practice

The "**Detached practitioner**" assessment practice has been touched on in the description of the "Detached practitioner" conception of assessment: their clinical knowledge and work were valued and prioritised over their assessment knowledge and work. This was both explicitly stated and observed through the approaching of assessment as a short-term, "today", administrative chore that simply needed to be done every clinical block. As these participants viewed practical task of assessment, negatively, they put little time, critical thought and effort into their assessment practice. They were largely passive in their roles and aimed for a hassle-free assessment practice, to minimise the impact on their more important clinical work.

"Our main job is not doing assessment, it is running the ward and doing other things, and it [assessment] is sometimes seen as a chore we have to do… It is very much a little extra thing that we do, it is not our main focus … [It is] something you just have to get over with, so it is not always the most effort is put into it … I think also, it would worry me that, not saying that I have given it more thought … [But] I am not sure about how it is working out there" (SA13).

(Negative perception/burden, little ownership/effort, concrete task, short-term perspective, clinical work prioritised, uncritical)

"I am a clinician first of all … We are all clinicians we are all doing this as extra, you know. So, I have a full clinical load … I have got other responsibilities" (SA16).

(Negative perception/burden, little ownership, clinical work prioritised)

"It is just because we are required to assess them as part of their mark" (SA13).

(Administrative, rules-based practice)

"I haven't actually thought about doing my own one [standardised rubric]" (SA13).

(Little ownership)

"I would assume [it] would be someone else's responsibility" (SA13).

(Little ownership)

A "Detached practitioner" assessment practice also involved confident decision-making based on clinical knowledge and expertise over the use of explicit, evidence-based, objective, standardised marking sheets. As will be discussed later, though, that basing clinical assessment-related decisions on clinical experience and expertise is not necessarily poor assessment practice.

"Someone else might run it differently, so there isn't really a standardised [assessment], even though the questions are standardised, the examination is often not standardised,

because you have different clinicians, different individuals who run the portfolio exam" (SA16).

(Little ownership, subjective practice)

"We have this proforma, this form, often we find it easier to just not sort of use it… We have *done it enough* to know what percentage to give, the rough mark to give the student and generally most of the examiners won't fill this in in detail, they will just give an overall mark … It's mainly just from our *own experiences*, when we were students, plus *getting to know the more you do*; it's easier to slot a student into the ranking of "excellent" to "barely okay" … Most people will just give an overall mark, which I suppose is *a feeling* that they get; how good the student is compared to the other ones that they have *examined before* … [Some examiners] are *quite good* at actually doing this *well… Every examiner may do it differently*" (SA13) – *emphasis* added.

(Clinical identity and disciplinary knowledge privileged, subjective practice, poor assessment literacy, craft knowledge)

"I think those sort of exams [clinical] are always going to be *subjective* … Most of the examiners … are also *just clinicians* who are working in the various hospitals … Most of them are *quite seasoned examiners* … Some examiners still don't like tick sheets, so, they'll like listen to the story the student is telling them and then give a global *impression* mark without ticking ... I don't even know if I think they are wrong, because you are so *experienced* with something that just sitting there like a robot and ticking, you know, just giving your *impression* on how you think the student managed the patient is maybe a better way of doing it, but I can't argue for or against either one because I think that is challenging, especially when you have someone who is so *experienced* and you ask them to tick box" (SA14) – *emphasis* added.

(Clinical identity and disciplinary knowledge privileged, subjective practice, poor assessment literacy, craft knowledge)

"[There is] no formal training programme for examiners unfortunately, [but] they all have standardised tools. The examiners are all clinicians … experienced clinicians, we are all experienced clinicians, but in terms of examination techniques we are all pretty much learning on the job" (SA16).

(Clinical identity and disciplinary knowledge privileged, subjective practice, poor assessment literacy, craft knowledge)

"One just develops skills over time, and, in assessing itself you see, you know what works and what doesn't" (SA17).

(Craft knowledge)

Another reason why these participants based their decisions on their clinical expertise was because they held a low assessment literacy and were thus dependent on others to direct their assessment practice. A rules-based assessment practice was adopted, whereby lecturers often passively and unquestioningly followed external rules and guidelines provided by their department or HEI. This was observed in practicing the traditional, historical or "status quo" practice of their department or HEI. An example of rules-based practice was directly implementing the university's policy on pass marks uncritically.

"It is just because we are *required* to assess them as part of their *mark …* It is what has *always been done* – which doesn't necessarily have to be the best way to do it, but, it is sort of the way we do it now" (SA13).

(Administrative, rules-based practice/status quo, little ownership)

"[We do] what we are expected to do in the faculty, they tell you, "This is what you are supposed to do"" (SA15).

(Administrative, rules-based practice/status quo)

"I think *historically* we have always had clinical assessments" (SA16).

(Administrative, rules-based practice/status quo)

The "**Detached practitioner**" assessment practice may be described as administrative, managerial, mechanistic and operational in a lecturer's approach to their conceived purpose and practice of assessment. These participants saw themselves as merely managing the physical or concrete assessment task on the ground, with little critical thought or effort being given into it, unquestioningly implementing routine practice. This was observed in their rules-based assessment practice where they followed the prescribed rules, policies and traditions of their department and HEI. As practitioners/clinicians, they based their assessment-related decisions primarily on their clinical experience and expertise (which is not necessarily poor assessment practice in HPE).

6.2.2. The Emerging equilibrium assessment practice

The most notable "**Emerging equilibrium**" assessment practice linked to the "Emerging equilibrium" conception of assessment was the belief that in order for these lecturers to be good teachers and assessors they also needed to be practicing clinicians. Being both a practitioner/clinician, a content and disciplinary expert, with personal experience with patients, *and* an educator, preferably with HPE training, was considered to be a requirement of academic staff at medical HEIs.

Unique to this category of assessment practice was its emerging nature. These participants expressed confusion of technical aspects of their assessment practice, for, as practitioners/clinicians, they were still developing their capability as educators and assessors. As they did not yet possess an advanced assessment literacy, and did not feel empowered to initiate changes themselves, they instead had to rely on others with HPE training or experience to guide them (this will be discussed in the HBT model of lecturer assessment practice in **Chapter 7**). What was distinct about this assessment practice was their questioning and willing nature, active and open, to change their assessment practice, in contrast to the "Detached practitioner" practice, where participants were content to privilege their clinical expertise above HPE.

"They all have standardised tools (and) the examiners are all clinicians… we are all experienced clinicians. But in terms of examination techniques, we are all pretty much learning on the job … I don't think we receive training on … education theory; because I am not up to date, I can't-, I don't feel empowered to actually change things, because I don't-, you know, that's not my expertise. So, I think a lot of the assessments we actually rely on (colleagues), people with an educational background to help us develop assessment tools … Does it work at the moment? I think our current assessment works. If you were to change it, I am not sure what we would change it to, I would need to know what the options are. I think currently we are assessing theory and I think we are assessing clinical abilities, which personally I think works. If there are other ways of doing it, I would be willing to explore those" (SA16).

(Questioning/uncertain, developing assessment literacy, lack of agency/requires help from others)

The "**Emerging equilibrium**" assessment practice reflected the conception of a dual identity. Both "Detached practitioner" and "Engaged educator" conceptions and practices of assessment were seen as important and necessary for achieving good assessment practice and outcomes in the "Emerging equilibrium" assessment practice. A mixture, of "Detached practitioner" and "Engaged educator" assessment practices were observed of participants in this middle category. These participants strove to achieve a balance of the two. Importantly, it could be argued, as displayed in the uncertainty of these participants in how to reconcile HPE evidence-based practice, of standard setting for example, and a reliance on clinical expertise for assessment-related judgements, that the identity of the Emerging equilibrium is possibly a transitioning identity, so, their assessment practice is still developing as they work towards finding clarity and confidence on their journey to "Engaged educator" and assessor.

6.2.3. The Engaged educator assessment practice

As seen in the description of the "**Engaged educator**" conception of assessment, these participants conceptually understood assessment to be an educational tool to assist students in their learning and development. They achieved this goal through the use of more continuous and longitudinal types of assessment, ungraded formative assessment tasks, reflection and the provision of feedback.

> "In their log books they have to do formative case presentations … and they get feedback from the consultants in terms on how they can improve and it doesn't count for a mark… It is purely for the student's benefit, there is no mark that goes along with it and the students are encouraged to use that as an opportunity to learn" (SA14).
>
> (Formative, student-centred)

> "The formative promotion of student learning, where the feedback helps to improve… The informal feedback … They [students] [are] welcomed as part of the health team, being given the responsibility to do procedures that they ordinarily wouldn't be able to do, feeling like they are functioning as a doctor than just a student in this big hierarchy … That's as valuable for their formation as … professionals to practice as the formal feedback on the report" (SA17).
>
> (Formative, long-term, competency)

These participants possessed a more advanced assessment literacy, and explicitly implemented sound technical assessment principles, using practices such as blueprinting and standard setting, where the numbers are not absolute but rather reflect competency.

> "I blueprint everything; so we have a curriculum, a core curriculum, that's our starting point … [That] outlines the learning outcomes and our students also have what we call a skills logbook. So whatever they are assessed on is directly matched to the logbook which is based on the learning outcomes which can be traced back to the core curriculum …

That's basically how we look at our assessments ... This whole thing about validity and reliability and all those things … [I] have the strong belief that you cannot just assess someone blindly … your starting point must be the core curriculum; they need to have this roadmap, "That is where you are going"; you need to have clear outlines of what you expect and you have to assess what they are exposed to" (SA15).

(Blueprinting, sophisticated assessment literacy)

"We have quite well-defined rubrics … There is a range, but we've got very clear descriptions for what fits within each range; so, a student who's clearly weak, doesn't know basic knowledge, can't make basic interpretations, deemed to be unsafe, that's sort of below 45% ... [If] they have assessed something incorrectly or they have clearly failed but you think it's redeemable, then they're 45% and so on and so forth" (SA18).

(Standard setting, sophisticated assessment literacy)

"The end of block exam is … an OSCE … it is very clinically orientated … We wouldn't ask them [students], "The 10 causes of this?" because that would be more rote learning, which is something that gets assessed earlier on; so it's more like, "This patient presents with this symptom or sign, how would you investigate it, this is the diagnosis, how would you treat it?" So it is more like a clinical scenario … it's a bit more interaction, so the clinical picture will evolve… The case presentation and the portfolio cases … the students are given a patient, a real-life patient to clark, to take a history from, to examine … [They] know how to speak to a patient, how to get a history, they need to know how to examine a patient … We are hoping that they are at a point where they know how to manage a patient, treat them properly … the scope of practice of an intern" (SA14).

(Educational impact, patient-centred, social and moral purpose, student development/competency)

Generally, while participants with "Detached practitioner" conceptions of assessment were unaware of the technical aspects of assessment, and employed a subjective, but robust, clinical expertise and intuition in service of assessment, and participants with "**Engaged educator**" conceptions had a high level of assessment literacy, participants from all three categories, despite their varied conceptions of assessment, performed a relatively similar and good quality assessment practice, in terms of using multiple assessment methods to achieve multiple purposes, with a strong clinical focus, in order to prepare students for future practice, whether or not they were themselves responsible or active in this assessment practice and its outcomes.

Claiming a causal relationship is beyond the scope of this study, for further investigation and validation are needed, and, a total alignment between said conceptions and practices was not always observed, for additional personal and contextual factors impact on lecturer assessment too (see **Chapter 7**).

6.3. Lecturers' Conceptions of Assessment and Practice in diverse Southern Contexts

A third phase of data collection and analysis took place within the confirmatory study, resulting in an additional 13 interviews with participants in another Southern setting (Mexico). Through Phenomenographic analysis, critical discussion and reflection, a third and final Outcome Space, encompassing and surpassing the previous two Outcome Spaces, was developed. The final Outcome Space consisted of four conceptions of assessment and eight more nuanced and rich dimensions, and represented the complexity of lecturers' conceptions of assessment (see **Table 6.4.** for an outline, for a table with illustrative quotes see **Appendix Table 6.2.**).

6.3.1. Final Outcome Space: Lecturers' Conceptions of Assessment and Practice in diverse Southern Contexts

In considering all the data collected, from three diverse Southern sites, four conceptions of assessment, held by lecturers in diverse Southern settings, were identified and labeled: "**Passive operator**", "**Awakening inquirer**", "**Active owner**" and "**Expert/Scholar**". Eight dimensions for each conception were identified: "purpose of assessment", "temporal perspective", "assessment literacy" "identity", "role and responsibility", "reflexivity", "accountability" and "emotional valence".

Dimensions present in the third Outcome Space, that overlapped with the previous two Outcome Spaces, are "purpose of assessment", "temporal perspective", "assessment literacy", "role" and "identity". These elements were either explicitly, or implicitly, present in the previous two Outcome Spaces, or expanded upon in the final Outcome Space. For example, "purpose", in addition to lecturers' conceptual understandings of assessment (concrete or abstract) and its purposes (administrative, formative, social and moral), also expanded to include the *scale* of their assessment practice (locally limiting it a single student or to their course, or considering it more globally in light of the entire programme, wider field of HPE or lifelong learning).

While new dimensions are present in the final Outcome Space, it was due to developing a more nuanced, complex and rich Outcome Space, further expanding upon the previous two Outcome Spaces, and reflecting the maturation of the researcher, in moving from comprehension, to synthesis, to theorisation, to coherently represent all the data collected (Morse, 1994). Within the Outcome Space itself, all the dimensions relate to at least another dimension, and categories were organised around to how dimensions sorted themselves within individuals and across all the transcripts. The titles of the categories themselves represent the hierarchical nature of the Phenomenographic Outcome Space and summarise the varied ways lecturers conceive of assessment.

To illustrate this, in the first Outcome Space, the ideas behind the "Undirected" category of assessment conception was one of poor assessment literacy, low reflexivity ("undirected"), a simplistic view of assessment as a short-term practical task they had to administer, an accountability to the HEI over any felt responsibility towards student learning or serving society, possessed and a negative attitude (emotional valence) towards assessment. All of these elements are present in the final Outcome Space. The categories of "Content-focused/Reproduction-directed" and "Competency and Conceptually-focused/Application-directed" are seen in the dimensions of the purpose of assessment, identity and role of the participant, and accountability. For the second Outcome Space, "identity" has now been placed as a dimension in the final Outcome Space. The labeling of the four conceptions has also evolved from "Detached practitioner" into "Passive operator", "Emerging equilibrium" into "Awakening inquirer", and "Engaged educator" into "Active owner", along with the addition the highest category of "Scholar". This increase in the number of dimensions and categories of description was not due to qualitative differences in the data sets, but, because of the deepening of the analysis and maturing of the researcher over time.

The dimensions of the "purpose" of assessment, which includes various elements of conceptual understanding, and "temporal perspective" relate to one another (they "sort together" within individuals). For instance, if a participant viewed assessment as administrative, that meant that typically they saw assessment as a concrete task and aimed at capturing student marks the administrative purposes of their course and hence resulted in superficial and short-term memorisation by students. In contrast, a participant who viewed assessment as serving a greater moral and social purpose, tended to see assessment as an abstract tool for student formative development towards competency for safe clinical practice of patients and the public, which takes place over a longer time frame, such as the entire programme, which revealed a long-term perspective. Similar relations exist between a participants' perceived role and responsibility, and their identity and sense of accountability. If a participant identified as a practitioner/clinician, they saw their assessment responsibilities as an administrative task and hence felt an accountability to the HEI. However, if a participant identified as an educator, then their role as

an assessor was similarly viewed as a guide for their students, desiring their learning and development, and hence felt accountable to their students and their development. As was seen in the second Outcome Space, identity is related to assessment literacy and reflexivity. Participants who identified as a practitioner/clinician, generally had a poor assessment literacy and were passive and uncritical in their assessment practice, revealed low reflexivity. On the other hand, participants who are developing ("emerging" or "awakening") reflect their emerging reflexivity as they asked questions or expressed uncertainty. Similarly, participants who seemed to identify as educators, and held a sophisticated assessment literacy, were active and reflexive in their assessment thinking and practice. Emotional valence has been present from the first Outcome Space.

Table 6.4.: Outline of the Final Phenomenographic Outcome Space reflecting of lecturers' conceptions of assessment in diverse Southern contexts.

<table>
<tr><th></th><th>Passive operator</th><th>Awakening inquirer</th><th>Active owner</th><th>Scholar</th></tr>
<tr><td rowspan="9">Purpose of assessment</td><td rowspan="2">Administrative (a task, to get marks)</td><td>Psychometric (to measure student learning)</td><td colspan="2">Psychometric
Moral & Social
(student learning, public service and patient care)</td></tr>
<tr><td colspan="2">Summative (marks for decision making)</td><td colspan="2"></td></tr>
<tr><td rowspan="2"></td><td rowspan="2">Formative (marks as feedback)</td><td rowspan="2">Formative (deliberate feedback for student development)</td><td rowspan="2">Formative (global feedback: student learning, teacher performance, assessment practice, curriculum appropriateness, practice preparedness)</td></tr>
<tr></tr>
<tr><td colspan="2">Concrete and Practical (a task)</td><td colspan="2"></td></tr>
<tr><td></td><td colspan="3">Abstract (a tool for conceptual development)</td></tr>
<tr><td>Administration (capture marks)</td><td>Content reproduction focused (academic/ knowledge-focus)</td><td>Competency development focused (knowledge, skills and attitude)</td><td>All stages of learning and development</td></tr>
<tr><td colspan="2">Local (individual student and course)</td><td>Global (student learning)</td><td>Global (student learning over a programme and life-time, considerations of HPE field more broadly)</td></tr>
<tr><td>Temporal Perspective</td><td colspan="2">Short-term (task, to get marks)</td><td>Long-term (to develop competency for future clinical practice)</td><td>Long-term (to develop competency for future clinical practice, lifelong learning and improve HPE practice)</td></tr>
<tr><td>Assessment Literacy</td><td colspan="2">Poor to developing assessment literacy (ranging from incorrect ideas, to disciplinary knowledge privileged, assessment</td><td>Developing and advancing (moving towards HPE</td><td>Sophisticated HPE knowledge</td></tr>
</table>

	experience/craft knowledge, to developing a basic awareness)		evidence-based understanding)	
Identity	Clinician	Clinician-educator	Education	Educationalist and researcher
Role and Responsibility	Administrator (manager, operator) and content-expert (disciplinary/clinical knowledge)	Assessor as a content-expert	Assessor as an educator (guide)	Scholar (contributes to educational research)
Reflexivity	Limited and passive (unquestioning or noncritical acceptance)	Active and questioning (becoming critical)	Critical and evidence-based practice (consumer of HPE, implement best practice principles)	Scholar (consumer and producer of HPE – researcher)
Accountability	To HEI (guided by rules and regulations, reputation/ standards)	To student (learning)		
	To profession and discipline (knowledge/content and reputation)		To patient and Public (clinical competency, moral and social)	
Emotional valence	Negative	Neutral	Positive	

The "**Passive operator**" participant believed that the primary purpose of assessment was summative and administrative in that it needed to provide a record of student's attendance and achievement in a course, almost a "tick the box" activity. This conception of assessment reflected a concrete conceptual understanding and a short-term and local view of assessment. This conception of assessment was found to be linked to a practitioner/clinician identity, because clinicians viewed their identity and clinical works as being of primary importance and which dominated any role and duty as an educator. This was observed in their reliance on clinical and disciplinary knowledge and a chief accountability towards their profession, using assessment as a summative tool to measure disciplinary knowledge reproduction and gatekeep who may enter and who may not.

As these participants did not possess any formal assessment or HPE training and held a limited assessment literacy and reflexivity, and a more negative emotional valence towards assessment (an added burden that they "had to do"). Interestingly, it appeared as though an accountability to the HEI sorted within this conception of assessment. This could be linked to the low reflexivity of these participants and their uncritical acceptance of the HEI's practices, what I call a "rules-based" practice, whereby these lecturers simply did what they were told to do by their department or HEI.

> "We are required to assess them as part of their mark" (SA13).
> (Administrative)

> "I would be very, very focused on trying to get my students pass any test" (MX9).
> (Administrative, summative, short-term, local, teacher-centred)

> "They need to pass that exam" (SA14).
> (Administrative, summative, short-term, local)

> "Educate a person … to take an exam" (MX4).
> (Administrative, summative, concrete task, short-term, local)

> "They come with me, and then you do like as a subjective evaluation, at the end you just check the list … Lots of the time it's very numerical … In the end, we have to assess students" (MX13).
> (Administrative, summative, concrete task, short-term, local, operational, uncritical/passive, accountable to HEI rules)

> "Evaluation still has to count towards a goal, it must still provide a mark" (SA1).
> (Administrative, summative, concrete task, short-term, local)

> "So, students come through, we just assess them and throw them out" (SA8).
> (Administrative, summative, concrete task, short-term, local, operational negative perception)

“A necessary evil” (SA6).

(Negative perception)

“I am a clinician first of all” (SA16).

(Clinician identity prioritised, operational assessment)

“I am a representative of the profession … I’m obviously protective over the [disciplinary] company, I can’t allow someone to go into the community, to sign off … and we let him go and then he kills people or he makes terrible errors in judgement … I mean that’s accusation against us … We have a tremendous responsibility towards … the university and our profession” (SA7).

(Clinician identity prioritised, accountable to HEI and profession)

“Our main job is not doing assessment, it is running the ward” (SA13).

(Clinician identity prioritised, negative perception)

“A chore we have to do” (SA13).

(Concrete task, administrative, short-term, HEI rules, no ownership/passive, negative perception)

“The University says it should be so” (SA6).

(Accountable to HEI rules)

“The university’s rules stipulate … We go about it the way we do purely because … the university recommends it” (SA1).

(Accountable to HEI rules, uncritical/passive)

“It is what has always been done, which doesn’t necessarily have to be the best way to do it, but, it is sort of the way we do it now” (SA13).

(Uncritical/poor reflexivity, rules-based, passive/operational)

> "It's a very subjective thing … I have to say it's a gut feeling … There is no objective or quantitative measuring instrument I use" (SA3).
>
> (Poor assessment literacy)

The "**Awakening inquirer**" conception of assessment shared some dimensions with the "Passive operator" category, however, it was more complex than the limited "Passive operator" category, because, a development ("awakening") from a basic knowledge and understanding, towards becoming more sophisticated assessment literacy was observed.

The "Awakening inquirer" conception of assessment centred around assessment as a psychometric and summative tool used to measure factual knowledge or content reproduction. These participants saw their role as more teacher-centred and content-provider. They identified as a practitioner/clinician-educator with their content expertise being based on their clinical experience. Similar to the "Passive operator" conception of assessment, the "Awakening inquirer" participant conceptually understood assessment to mostly be a concrete, short-term and local task performed to generate marks for the students in their clinical rotations. Yet, as they saw themselves as both a practitioner/clinician *and* educator, assessment was also seen as a potential tool for student learning, although the goal of that learning was information regurgitation.

The developmental nature of this category was displayed in their emerging reflexivity, observed in their questioning and critique of themselves and their assessment practice. They were beginning to admit the limitations of their technical knowledge of assessment and reliance on their clinical expertise over HPE (which is also linked to their joint clinician-educator identity). These participants were starting to see the value of HPE and assessment training, which perhaps triggered this awakening awareness and more active thinking about assessment, even if they did not yet possess the assessment literacy or abilities needed to enact a high-quality assessment practice.

"[Testing the] minimum level of factual knowledge" (SA3).

(Content-reproduction, short-term, psychometric/summative)

"Make sure that there is a certain standard of knowledge or certain level of knowledge" (SA16).

(Content-reproduction, short-term, psychometric/summative)

"I'm going to ask you this therefore you must learn it" (SA6).

(Teacher-centred, content-reproduction, psychometric/summative)

"What the lecturer thinks is important" (SA12).

(Teacher-centred, teacher as content expert)

"The professor … he's the expert, the teacher has to give the information" (MX4).

(Teacher-centred, teacher as content expert)

"A clinician trying to be a teacher" (MX8).

(Dual identities)

"I am a clinician convinced that the way I practice needs to have education in sight… Finding that halfway thing is something I believe" (MX3).

(Dual identities)

"The role is mixed, balanced clinician and teacher – equal" (MX6).

(Dual identities, equally important, balance/equilibrium)

"We have like different hats in different moments in the hospital, because our classrooms are inside the hospital and, let's call it … side-by-side with the school, the hospital. So it is many roles at the same time … So, the identity is like mixed … We don't want professor that only has like theoretical knowledge and have never seen a patient … Our professor cannot teach something he had never done … We need a balance between those [the

teacher and the clinician] … We need to have a balance between those two peoples. So, I am in the middle of those [valuing teaching and clinical work] … That is very important because if you don't have both points of view in common you cannot do all of this" (MX1).
(Dual identities, balance/equilibrium)

"For most doctors it happens [prioritising the clinical over the educational work]. I don't give that [clinical work] much importance; I mean, for me, they are the same the level of importance, but for most doctors in this course, private practice [clinical work] is more important than the clinical course [educational work]. If they have to do something and they are not able to go to the class they cancel the class and they go see their patient … I believe both [clinician and educator]; yes, both. I mean I do like the teaching part, so both … Also, I believe in medicine you can't be a teacher if you don't have enough experience clinically" (MX2).
(Dual identities/balance)

"I also realized is that I actually had very little idea of what to do … We actually had lectures about lecturers … that's how I got involved, and I think because one starts to ask questions … The first time we started thinking about assessment … we learned about different assessment methods and so there I started thinking … It had value to me, but also frightened me a bit, in the sense that I knew I was lacking" (SA10).
(Awareness/reflexive, developing assessment literacy)

"I hope we don't make mistakes that often, but I have a sinking feeling that we do make them quite often … We make mistakes" (SA2).
(Awareness/reflexive, developing assessment literacy)

> "We often assess incorrectly and because we don't ask higher order questions, we don't test deep knowledge, but rather we test surface knowledge which they can memorise very easily and can regurgitate" (SA4).
>
> (Awareness/reflexive, developing assessment literacy)

The "**Active owner**" conception, like the "Awakening inquirer" conception, shared dimensions with both the preceding and subsequent conceptions, depicting the relational and hierarchical (potentially transitioning and advancing) nature of the Outcome Space.

The more sophisticated nature of this conception was seen in the more complex and nuanced knowledge and understanding of assessment. For example, assessment was not only believed to be a summative tool for decision-making but also a formative tool to guide student learning. These participants conceptually understood assessment to be an abstract tool for student conceptual growth and competency development as future health professionals. This was observed through their assessment practices that focused on the application of knowledge in clinical settings, which also revealed a more global and long-term conception of assessment.

As competency development was the goal of assessment, these participants saw themselves as educators and guides, and felt accountable to their students, their learning and professional development. To this end, they practiced formative assessments and gave feedback to their students. Assessment was seen as important and valuable because there was a social and moral function in preparing future practitioners for clinical practice and patient-care. This value was also observed in their positive emotional valence towards their assessment practice, for, if they perceived assessment as important and valuable, then they desired to use assessment to inspire learning and help students grow. This was related to their educator identity, role and responsibility, because they taught due to a personal interest in and enjoyment of education.

Participants with an "Active owner" conception of assessment possessed an advanced assessment literacy. This was displayed through their use of terms such as "rubrics" (referring to

structured and more objective measures of grading) or "mapping" (related to milestones and blueprinting). Related to their identity and assessment literacy was the fact that these participants often had experienced basic HPE training. These participants understood that assessment served multiple purposes and practiced varied assessment methods to ensure more reliable and valid assessment of their students. While these participants were not HPE experts or scholars, they had moved on from an implicit craft knowledge of relying on clinical experience and expertise alone, to implementing HPE evidence-based principles. This category of participant was more in control and had an active ownership of their assessment practice, which was in line with their possession of some HPE knowledge and increased agency.

> "I believe it [assessment] is a tool to see if the student gets prepared or gets ready to do the objectives that we want him to have when he becomes a doctor" (MX2).
> (Abstract tool, formative, competency, social and moral, long-term)

> "Are they able to connect the classroom knowledge to the real-life setting?" (MX13).
> (Application of knowledge/competency, patient-centred, global, long-term)

> "Whether or not he [a student] is informed enough for… practice one day" (SA5).
> (Abstract tool, formative, competency, social and moral, global, long-term)

> "Competence, competence, competence" (MX6).
> (Abstract tool, formative, competency, global, long-term)

> "Your role will be to advise students and support them" (SA1).
> (Facilitator/guide, educator identity)

> "As a professor I can say, "Oh, it's your fault, you're not studying, it's your fault" but it's my responsibility we have to share … It's a shared responsibility between students and teachers … we are in the same boat" (MX4).
> (Reflexive, accountable to student)

“That constant dialogue with a student is a feedback, it is not a grade, it is a feedback” (MX4).

(Formative)

“Feedback … how they can improve … It is purely for the student’s benefit … an opportunity to learn” (SA14).

(Formative, student-centred)

“You try to identify in yourself which areas are you weak, that you are going to learn about … that you are going to take action in” (MX11).

(Formative)

“The development of the student as they grow … progress … improve” (MX1).

(Formative, competency, global, long-term)

“[Our] formative assessments … We have to help them; we have to do something with them … And it usually results in a massive improvement” (SA7).

(Formative, accountable to student learning)

“Assessment, in the end, has to be a reflection of … how the student has grown in terms of that subject field, how he has developed” (SA4).

(Formative, competency, global, long-term, accountable to student learning)

“A way to measure the development of the student as they grow through the different steps … The student can see their own progress … We are assessing the development of the student … If they have strong capacities or weak capacities, so they can identify it and improve it … I do the assessment every day … it is in real-time. That’s important because at the end of the rotation they [students] can see how they did at the beginning of the month and the goal is that they improve … their abilities along the rotation, so, at the end

of the month, it's supposed to have high points in every assessment … We can see the improvement of the students, first week, second week, third week and also the fourth week … Every week we see how the students are doing … they have a progressive score; so we can see which students get the same points or the students that have a problem or the students that are doing well in the clinical areas"" (MX1).

(Abstract tool, formative, competency, global, long-term)

"To become a better medical student and a better physician" (MX3).

(Formative, global, long-term, social and moral)

"Long term it has to do with the certifying someone as competent to a particular task – the responsibility towards society" (SA12).

(Social and moral, accountable to patient/public)

"Getting students to acquire the knowledge that they need to practice with … And I suppose in that sense it would safeguard society" (SA11).

(Competency, social and moral, accountable to patient/public)

"Responsible to civil society … good practitioners" (SA8).

(Social and moral, accountable to patient/public)

"We have [a] strong assessment, our quality control is high … [it's] a moral thing" (MX1).

(Social and moral, accountable to patient/public)

"I do not give them a final exam, I don't believe in final exams, I think they [are] not good… I think final exams are just a test of how much you can cram in your head … It is not even real life for a physician – every day is a test … What we should be trying to evaluate is how I'm able to solve that problem … Because it prepares you for real life. I believe in multiple assessment every day" (MX9).

(Formative, global, long-term, sophisticated assessment literacy)

“You should be assessed by multiple people … You evaluate every single thing, and I think it's actually more … fair for the students … I believe in in multiple assessment and doing it a lot of times … I use a rubric” (MX9).

(Sophisticated assessment literacy)

“That mapping part I was talking earlier, when the student knows which is the specific learning objectives for each learning opportunity ... So, the student starts reading the map … Those are the actions that are to be carried out by the student to achieve those particular learning objectives ... The student follows those things to get all the way to learn that learning objective” (MX3).

(Blueprinting, sophisticated assessment literacy)

“I have to say honestly that until a year ago I wasn’t a big advocate of the whole story of continuous evaluation … I believed in giving a man a proper exam, forcing him to bite his nails and sit on his backside for a day or two to go through the work. Unfortunately, that is also how I studied … *I have now developed other insights* … I have come to be a big advocate of the whole concept of continuous evaluation” (SA1) – *emphasis* added.

(Reflexivity)

“I do think that I enjoy (assessment); so no, I don't take it like as a negative” (MX13).

(Positive perception)

“I enjoy it [assessment]” (MX9).

(Positive perception)

“What I find wonderful about [our assessment], that’s really positive for me, is that they look at how we can help people” (SA7).

(Positive perception, accountable to student learning)

> "Inspiring people to learn" (SA1).
>
> (Positive perception, accountable to student learning)

> "It's my own interest … I was really involved in medical education … I was fascinated by the outcomes, and curriculum design is one of my favourite parts … I was really fascinated with the idea about this outcome evaluation process, how we could really do it … to see really what is going on with this student is fascinating; how the portfolio could help us to do that better. So I like evaluation, that's why I was involved in it and I try to keep involved" (MX12).
>
> (Positive perception)

The findings supporting a further conception, here called "**Scholar**", were less robust than for other categories. There were, however, clear indications of a conception beyond "Active owner". What distinguishes this category from the "Active owner" category is mainly the scholarly nature of this conception. Assessment was no longer related to a single student or clinical rotation but situated within the larger academic programme *and* field of HPE, indicating a global scale. For example, both the "Active owner" and "Scholar" conceptions shared the dimension of assessment purpose as being "formative", yet, the "Active owner" conception understood assessment as a tool to shape a student's learning, while the "Scholar" conception saw assessment as a two-way street, not only providing feedback to students to guide their development, but also providing the participant with important information to improve their own assessment practice, or to make changes to the larger curriculum and programme, or even reveal an area of potential HPE research. This was linked to an educationalist and researcher identity, for they not only consumed but also sought to contribute to HPE.

Similarly, the "long-term" temporal perspective dimension was shared between the "Active owner" and "Scholar" categories, however, while the "Active owner" conception was concerned with competency, the "Scholar" category took the conception further to include life-long learning, and, again, of both the student and the lecturer.

Additionally, another important distinction from the "Active owner" category was that the "Scholar" participants was indeed an expert in HPE, as was seen in their explicit assessment literacy and implementation of HPE evidence-based practice, but also their desire to do research and further the HPE field. These participants were reflexive and held qualifications in HPE, which contributed to their assessment literacy, identity, role, responsibility and agency, and they were confidence in their initiation and implementation of new, innovative and evidence-based assessment practices.

"Assessment goes both ways – with the professors and students" (MX7).
(Mutual accountability and learning, reflexive, global)

"I do give them a lot of feedback, and I expect feedback from them" (MX13).
(Mutual accountability and learning, reflexive, global)

"We do review our programme every year … we look at student feedback, which I think it important, and we try to improve" (SA15).
(Global, reflexive, HPE-based)

"Assessment is going to try drive learning, but it also drives teaching … The assessments are the most important thing in medical education" (MX6).
(Global, reflexive, HPE-based, positive perception)

"There's like this research team that looks for the new tendencies of other schools who try to imitate or listen to them … schools have their own like self-assessment to see what works what was good or what was wrong" (MX6).
(Consumer and producer of HPE, educationalist identity)

"I can validate because there is evidence that what we need to do something [new]. Sometimes we have been like the first pioneers in these kind of fields in Mexico, but there is evidence that we need to go over there [new fields of HPE]" (MX12).

(Global, reflexive, HPE-based)

"Programmatic assessment is our goal … Programmatic assessment as a guarantee, because of cycles of measuring, reflection, giving thought to the results and then making a change, like strategies" (MX6).

(Sophisticated assessment literacy/HPE-based)

"I think students should be responsible for their learning, and I'll say this, I'm thinking of me right now, as I'm 50 years old, and I'm still learning" (MX9).

(Global/mutual lifelong learning, reflexive)

"I'm going to give you [students] the tools that you are able to evolve in your learning of medical practice – and that's done through a lifetime. If you stop learning you better quit" (MX13).

(Global/mutual lifelong learning, reflexive)

To summarise, upon conducting further interviews and reviewing all the data collected in its entirety, a third and final Outcome Space was developed. It consisted of four categories of conceptions of assessment: "**Passive operator**", "**Awakening inquirer**", "**Active owner**" and "**Scholar**", and eight dimensions. The hierarchical characteristic of Phenomenographic Outcome Spaces was clear in the third and final Outcome Space, for it represented a continuum of lecturers' conceptions of assessment, ranging from basic, passive, uncritical and operational, to awakening and developing, to active, advancing, sophisticated and responsible, to assessment expert, researcher and scholar.

6.3.2. Southern Conceptions of Assessment: Summary of Phenomenographic findings

The third phase of this research project produced a third and final Outcome Space, which described four conceptions of assessment. To support these conceptions, archetypes for each category are displayed below: "Passive Operator" as **SA1** and **SA13**, "Awakening inquirer" as **SA3** and **MX4**, "Active owner" as **SA4** and **MX6**, and Expert/Scholar as **MX6**, **MX12** and **MX13** (see **Table 6.5.**).

Table 6.5.: Final Outcome Space: Archetype of lecturers' conceptions of assessment in diverse Southern contexts.

Passive operator (SA1 and SA13)	**Awakening inquirer (SA3 and MX4)**	**Active owner (SA4 and MX6)**	**Scholar (MX6, MX12 and MX13)**
"A chore we have to do" (SA1) *"Evaluation is a way of proving that the students have gone through the work" (SA13)* *"They have to know their stuff by the following Friday … We are required to assess them as part of their mark" (SA13)* *"Evaluation still has to count*	*"[Testing the] minimum level of factual knowledge" (SA3)* *"Educate a person … to take an exam" (MX4)* *"I'm a professor, but I am a clinician" (MX4)* *"The professor … he's the expert, the teacher has to give the information" (MX4)*	*"Competence, competence, competence" (MX6)* *"(This HEI) does not prepare students to pass an exam; it's for them to be good practitioners and to be competent" (MX6)* *"Insight … deeper level … [Assessment] has to actually be in line to help students improve their skills or their abilities to become better doctors" (SA4)*	*"Assessment is going to try drive learning, but it also drives teaching … The assessments are the most important thing in medical education" (MX6)* *"Are they able to connect the classroom knowledge to the real-life setting?" (MX13)* *"I'm more comfortable with saying that he's "doing okay", he's "proficient" or he's "lacking" – I like that way better" (MX13)* *"There are scales on "great performance", "he could do that better", "he needs to improve a lot", "he failed"" (MX12)* *"I understand the portfolio may be something more robust … At the end of the clinical rotation, the students have about 30 evaluation forms of different aspects with different ratings*

towards a goal, it must still provide a mark" (SA13) *"Our main job is not doing assessment, it is running the ward" (SA13)* *"Our own [clinical] experiences" (SA13)* *"It is what has always been done, which doesn't necessarily have to be the best way to do it, but, it is sort of the way we do it now" (SA1)*	*"I like to think that by asking different students different things, I can gauge an overall impression of their place on the marks scale. However, I know it's certainly not something one can prove scientifically … It's a very subjective thing … I have to say it's a gut feeling … There is no objective or quantitative measuring instrument I use" (SA3)*	*"Assessment, in the end, has to be a reflection of … how the student has grown in terms of that subject field; how he has developed" (SA4)* *"Seven competencies … So there are things that the student needs to learn how to do, professional activities, or to have these experiences" (MX6)* *"The first thing is that you'll have to define what the characteristics of … what the level of knowledge is that one would expect of a first-year student, on that level, what is the level of knowledge of a second-year student, and what level of knowledge does a third-year*	*… (The) portfolio gets everything; and I can have a better picture of my students" (MX12)* *"I do give them a lot of feedback, and I expect feedback from them" (MX13)* *"The challenge is how to help the students (learn) and how to help the professors to teach [and assess]" (MX6)* *"I'm going to give you the tools that you are able to evolve in your learning of medical practice. And that's done through a lifetime. If you stop learning you better quit" (MX13)* *"It's my own interest … I was really involved in medical education n… I was fascinated by the outcomes, and curriculum design is one of my favourite parts … I was really fascinated with the idea about this outcome evaluation process, how we could really do it … to see really what is going on with this student is fascinating; how the portfolio could help us to do that better. So I like evaluation; that's why I was involved in it and I try to keep involved" (MX12)* *"I do think that I enjoy [assessment], so, no, I don't take it like as a negative" (MX13)* *"There's like this research team that looks for the new tendencies of other schools who try to imitate or listen to them … schools have their own like self-assessment to see what works what was good or what was wrong" (MX6)*

		student have, so one can determine a profile for each and then … you can set up your questions around that profile" (SA4)	*"There is evidence that what we need to do something [new]. Sometimes we have been like the first pioneers in these kind of fields in Mexico; but there is evidence that we need to go over there" (MX12)* *"Programmatic assessment is our goal … Programmatic assessment as a guarantee, because of cycles of measuring, reflection, giving thought to the results and then making a change, like strategies" (MX6)*

In support of the findings from the first and second phases of this study, in the third phase of this study, it appeared as though there was a relationship between conceptions and practice too.

6.3.3. Lecturer Assessment Practice in diverse Southern contexts

The categories of the final Outcome Space relate to participants' approaches to their practice of assessment. The same category titles were used, with slightly different "dimensions", to describe the link between the espoused and enacted conceptions (see **Table 6.6.**). For example, how participants practiced assessment, their assessment outcomes and their role as an assessor, were linked to conception dimensions of assessment purpose, assessment literacy, temporal perspective, identity, role and responsibility, and accountability.

For example, participants with a "**Passive operator**" conception of assessment viewed assessment as a burdensome administrative task that they were *required* to perform (linked to a "rules-based" practice), believing that their time would be better spent on their more important clinical work, which led to the goal of practicing the most convenient assessment method as possible (for instance, what was "easiest to mark" and linked to poor assessment literacy). They used assessment as a local, short-term and summative activity to assign marks to students for a single course. Since these participants saw themselves primarily as practitioners/clinicians, and

their assessor role as administrative and operational, their approach to assessment was uncritical (linked to low reflexivity and responsibility), and passively followed HEI rules and regulations and the traditional assessment practice of their department. Participants holding these conceptions believed that their assessment practice resulted in superficial, short-term learning and a mark-focus by students.

The reverse was then true for the "**Active owner**" assessment practice, in which various assessment methods were used in order to achieve the multiple purposes of assessment (linked to conceptual understandings of summative, formative, social and moral functions), as well as their implementation of HPE evidence-based practice, both of which revealed the advanced assessment literacy of these participants. As these participants identified as educators, they felt responsible to the student for their learning and, so, practiced feedback to assist in their students' development towards competency for future clinical practice (linked to long-term perspective, role and accountability).

Table 6.6.: Lecturer Assessment Practice in diverse Southern contexts.

	Passive operator	**Awakening inquirer**	**Active owner**	**Scholar**
Assessment design	Limited, directed by others, for example HEI rules and requirements	To generate valid and reliable marks	Thoughtful, methods selected for format to fit function Information gathered for extrapolation to patient care	HPE evidence-based
Function of assessment practice (goal)	Administrative (HEI requirement)	Psychometric (measurement)	Moral and Social (student learning, public service and patient care)	
	Summative (decision-making, HEI requirement)	Summative (decision-making for a course)	Summative (decision-making for certification - competency for clinical practice)	

			Formative (student learning and competency development - transformative)	Formative (student learning and competency development, teacher performance, curriculum appropriateness = global practice)
Level of assessment practice	Concerned with meeting requirements outlined for courses by HEI		Engages with what assessment data means beyond the course	
Focus of assessment practice	Follows the rules of the HEI	Follows the tenets of the discipline	Does what students need to learn the most	Comprehensive practice engaging all forms of learning
Assessment Practice	Traditional and intuitive (follow HEI requirements and status quo unquestioningly), convenience	Traditional but starting to question	Critical and evidence-based (varied tools/methods used for different purposes)	Evidence-based (continuous assessment, feedback), Adaptive expertise (able to apply HPE knowledge and skills to different settings), creative and innovative
Perceptions of learning effects of assessment	Poor/superficial learning	Superficial learning (memorisation and short-term recall)	Deeper learning (study to understand, applied for clinical practice)	Deeper learning and clinical competency
Assessor Role	Assessor as administrator (practical manager, operator, mechanistic)	Assessor as judge and gatekeeper (decision-making and protective of profession/reputation)	Assessor as a guide (gives feedback) and gatekeeper (competency certification)	Assessor as a scholar/ researcher (critical consumer and producer of HPE best practice), innovative

Responsibility and ownership (priority)	Not responsible (merely does what told, little effort given), prioritises clinical work	Balance/compromise between clinical work and assessment (gives time to both)	Takes responsibility/ active ownership, reflects on outcomes of assessment, adapts assessment practice	HPE prioritised, critically engages with and practices HPE-evidence based assessment

While I propose that the findings from this study suggest a potentially causal relationship between lecturers' conceptions of assessment and their assessment practice, conceptions alone do not entirely account for lecturer assessment behaviour. As outlined in **Section 3.3.1.** and **3.3.2.**, numerous studies have shown that while conceptual change is needed in order to change assessment practice, changing conceptions alone is not sufficient in leading to altered and sustained practice. Indeed, disconnects and mismatches between conceptions and practice have been observed. For example, lecturers possessing a more sophisticated conception of assessment, yet practicing a poor assessment practice due to resource constraints, or lecturers with a simplistic and limited conception of assessment practicing a seemingly advanced and valid assessment practice due to blindly implementing good HEI assessment policies and rules. Using HBT as a guiding framework, additional personal and contextual that impact on lecturer assessment practice in diverse Southern settings was explored.

Chapter 7: Factors influencing Lecturer Assessment Practice in Southern contexts

> "I think that illustrates largely [the] really bizarre dichotomy between what you [would] like to be assessing and what happens in reality" (SA6).

Lecturers conceptions of assessment were described in **Chapter's 5** and **6**, and, while conceptions have been shown to be an important factor in shaping lecturer assessment practice (a possible link between espoused and enacted conceptions detailed in **Sections 5.3.**, **6.3.** & **6.4.3.**), additional personal and contextual factors, as purported in HBT, were also found to have a significant impact on how assessment was practiced.

In this study, **personal factors** influencing assessment practice included perceived agency[9], training and/or formal education in HPE, value, motivation and attitude, which were related to internal cost and response appraisal, and age. **Contextual factors** influencing assessment practice included perceived barriers, and related external costs, interpersonal, departmental, institutional, national and political factors (see **Figure 7.1.**)

While personal and contextual factors identified were separated into different sections in this thesis, in reality, and in line with HBT, they interacted with one another and were deeply entwined in their impact on lecturer assessment practice. Indeed, it was difficult to separate out single factors, for, they related to many other factors. However, personal factors are presented first, followed by contextual factors.

[9] Agency refers to an individual's capacity or power to do something, such as make their own choices or decisions, and is related to self-efficacy, which refers to an individuals' belief in their ability to achieve something or to execute or perform a behaviour.

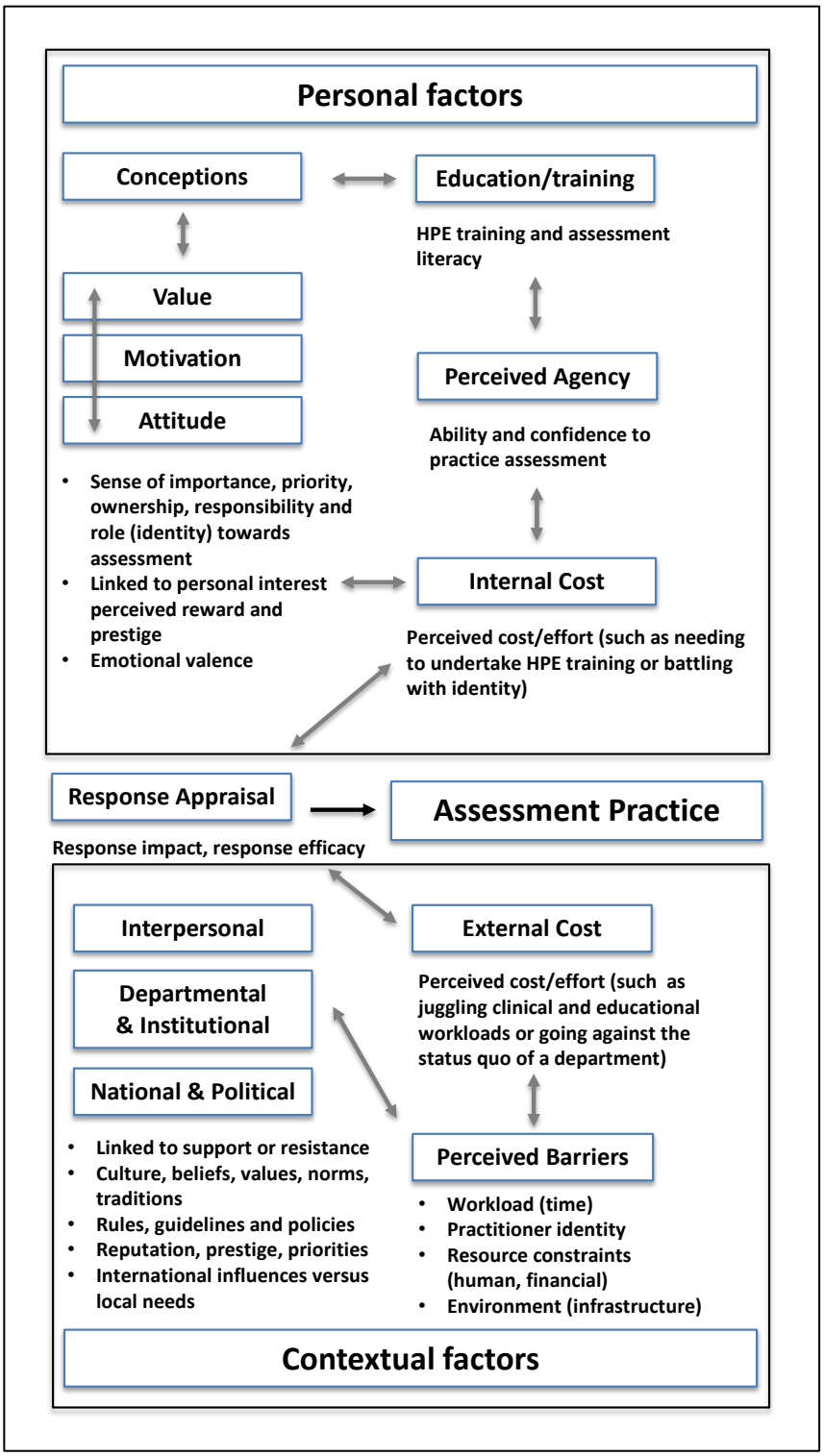


Figure 7.1.: A general model of lecturer assessment practice in diverse Southern contexts. A number of personal and contextual factors have been shown to influence lecturer assessment behaviour, including a lack of HPE knowledge or training, lecturer identity, a lack of time or resources, programme structure, collegial support and HEI culture.

7.1. Personal factors influencing lecturer assessment practice in diverse Southern contexts

The following **personal factors** impacted on assessment practice: conceptions of assessment, and their related dimensions (assessment literacy, identity, role and responsibility, reflexivity, accountability, emotional valence), perceived agency, education and training, value, motivation, attitude, response appraisal (response impact, response efficacy and cost)[10], and age. These factors, along with contextual factors, interacted with one another, to shape lecturer assessment behaviour.

In this example, a participant displayed a low assessment literacy, and a related low perceived agency, due to a lack of formal assessment training, using their practitioner identity as an excuse. This lead to a reliance on job learning and craft knowledge development, over HPE evidence-based practice. Yet, due to a developing reflexivity, seen in their uncertainty and questioning, acknowledging their lack of confidence, and a need for formal HPE, they turned to colleagues with educational training to guide them. This suggested a growing value and appreciation for HPE, however, at this stage, this individual lacked the ability to direct quality changes themselves and had not sought further assessment training. Importantly, relying on an educationally experienced colleague could remove the felt need of this individual to seek HPE training and take on more responsibility and ownership of their assessment practice, allowing them to continue prioritising their more valuable clinical identity and work over HPE.

> "[There is] no formal training programme for examiners unfortunately … In terms of examination techniques, we are all pretty much learning on the job … I don't think we

[10] Response appraisal refers to how a participant appraises a potential response or action taken. For example, what would the effect of this response (assessment behaviour) be? Will it bring about the effects I seek? Which includes response efficacy: how much effort would this require of me? Is the potential benefit high or low? And, the internal and external response costs: what are the perceived demands of this action taken, costs to myself, and are they worth it. Response appraisal also ties in with value, motivation, attitude, identity, and perceived barriers, which relates to response impact: what would the (size and likelihood of) impact be?

receive training on how to actually set a good, valid MCQ-type question … I think a lot of the education theory, because I am not up to date, I can't-, I don't feel empowered to actually change things, because I don't-, you know, that's not my expertise. So, I think a lot of the assessments we actually rely on [colleagues], people with an educational background to help us develop assessment tools … Most departments will have a course convenor [this individual], and they then report to the year convenor, who reports to the programme convenor. So [names colleague] is like an extra level, which is quite nice. So, [names colleague] will oversee all the undergraduate teaching [and assessment] … I'm a clinician, I am not [a teacher], I haven't got any experience in education and my main interest is clinical work … So, a lot of our direction comes from [names colleague] because [names colleague] has got postgraduate experience and training in education … [Names colleague] actually designed our portfolio mark sheet" (SA14).

(Lack of HPE training leading to low assessment literacy and low agency, and a reliance on supportive colleagues with educational training, need and value of HPE training realised, developing reflexivity, yet a lack of ownership and privileging of practitioner identity and work)

Similarly, many participants found assessment to be difficult, revealing a low agency and high internal cost, and which led them to operational and rules-based assessment practices, seen, in this example, by an implementation of the most recent assessment platform provided by their institution. This participant did not understand what they were doing, or why, and simply used the assessment platform provided by their HEI, believing it to be good enough. The fact that they were unable to design and initiate a new assessment practice themselves, also showed a lack of assessment literacy, reflexivity and ownership.

"I think it's [assessment] the most difficult things that I have to deal with … And, as I told you, I don't have the answer. I don't know. I've tried. We have jumped over models, different models, to see what it works … We jumped to a platform from our [HEI] … So it's been hard … It's very difficult to make an assessment" (MX7).

(High internal cost/low agency, leading to a passive HEI/rules-based practice)

Other participants acknowledged their limited assessment literacy and agency, because of an exposure to HPE. HPE enlightened participants to their deficits, and awakened their reflexivity. In light of these disorientating (transformative) experiences, participants became more active in their assessment practice. Some experienced an increase in their assessment literacy and agency, others grew to value assessment more. This led one participant to change their assessment practice, through asking questions in a different manner, and, another participant changed their assessment practice from high stakes to continuous. However, when some participants attempted to introduce new HPE evidence-based assessment practice in their departments, they experienced resistance from colleagues in their departments, because those practitioners valued their clinical identities and disciplinary knowledge above HPE. This provided a barrier to changing their assessment practice.

> "I think my knowledge in that regard is a bit *too limited* to tell you about the educational purpose [of assessment] … Once again, my knowledge of purely educational concepts is *not so extensive* … We're still doctors who teach a few classes … And the other thing I also realized is that I actually had *very little idea* of what to do … We actually had lectures about lecturers … It had value to me, but also frightened me a bit, in the sense that I knew I was lacking … The first time we started thinking about assessment was when we had the new course … [We] sat and we learned about different assessment methods, and, so there, I started thinking … [But] it didn't necessarily go through to the rest of the department … I don't think there's necessarily always an opportunity for my opinions and the opinions of my colleagues in the department … [Because] if a guy says he's a specialist in a [clinical] field, [then] there's nothing you can teach him, "We've done it like that for years, so why would we change it?"" (SA10).
>
> (Exposure to HPE revealed low assessment literacy and low agency, reflexivity activated, yet practitioner identity/disciplinary knowledge privileged, relating to departmental culture/normative beliefs/values, tension between supportive and resistance colleagues)

> "The other thing is that teachers like us, we are doctors, we are not teachers, nobody taught us how to teach, and nobody taught us how to do [design] a test … Somebody

needs to teach us how to teach how to perform assessments in different ways … I set the questions and someone, an educationalist … [gave] comments, and there were points, maybe small things, but when I changed them I said, "This is different, it makes sense, the main information is the same maybe, but it's put in a different way – it has a lot of sense now." So if I didn't do that course, I would be doing the same thing … because nobody told me, "That's wrong, your test is wrong, I'm sorry" – teach me!" (MX8).

(Exposure to HPE training led to a changed assessment practice)

"I have to say honestly that until a year ago I wasn't a big advocate of the whole story of continuous evaluation … I believed in giving a man a proper exam, forcing him to bite his nails and sit on his backside for a day or two to go through the work. *Unfortunately that is also how I studied* ... But, *I have now developed other insights …* I think one has to move towards continuous evaluation slowly but surely, even if it means small tests about smaller parts of the work … It's the whole story of the new curriculum" (SA1) – *emphasis* added.

(Exposure to HPE/new curriculum led to a changed assessment practice)

"I can't remember where, but we actually had *lectures about lecturers* … It had *value* to me, but also *frightened* me a bit, in the sense that I knew *I was lacking*. I decided to be *serious* about it and even came to [an HPE expert] once or twice and for my assessment [practice] … you know, that whole issue of fairness … So that's how I got involved [in HPE], and I think because one starts to ask questions … I think people were so glad that one guy in the Department does that" (SA10).

(HPE training, leading to increased value, reflexivity, assessment literacy and ownership in their assessment practice)

On the other hand, some participants were unaware of their lack of assessment literacy (low reflexivity), and confidently practiced a subjective assessment practice, as opposed to standard setting, based on their assessment craft knowledge. This too revealed a privileging of their clinical identity, experience and expertise over HPE.

"It's a *gut feeling* one just has, one that *I've picked up along the line*, you know, and I've realised that when I set a paper and I sit and look at it, and I tell myself, "This is more difficult than usual." Then you'll see a correlation between what I feel and the marks we end up getting. So, it's something that comes with *experience and time*, you get to know where the standard lies, but *there isn't an instrument I can measure it with*, hold it up to a benchmark and say, "These are the standards I have to strive towards"" (SA2) – *emphasis* added.

(Practitioner identity and experience/craft knowledge privileged, leading to a subjective assessment practice)

"I am aware that some people who know about training and assessment feel that even an oral has to be structured, rather than it being just a subjective type of thing from the examiners side. *I am not convinced* about that at heart. Maybe I have an *inflated idea of my own capacity* to, to evaluate students, but I like to think that by asking different students different things, I can gauge an *overall impression* of their place on the marks scale. However, I know it's certainly not something one can prove scientifically … It's a *very subjective* thing … I have to admit... And it's a thing that *comes with experience*. For example, the first few years I didn't know what to do [but] now I know from … [what] we've done in previous years … And it's basically about that. So, I have to say it's a *gut feeling*. There is *no objective* or quantitative measuring instrument I use" (SA3) – *emphasis* added.

(Practitioner identity and experience/craft knowledge privileged, low assessment literacy, leading to subjective assessment practice)

The personal factor of motivation was found to be related to personal interest and perceived reward (inherent motivation), value and the attitude. For example, participants who had a negative attitude towards assessment, often due to perceived barriers, such as the competing demands of clinical and assessment workloads, also had a low motivation and high costs towards assessment (influencing response appraisal, such as choosing to not invest in their assessment

practice). This, too, was often linked to a practitioner identity, which meant that clinical workload was more highly valued and prioritised. These factors resulted in less time and effort being put into the assessment practices of these participants.

> "Our main job is not doing assessment, it is running the ward and doing other things, and it [assessment] is sometimes seen as a chore we have to do, especially if it is every two weeks. So, it is … something you just have to get over with, so it is not always the most effort is put into it … [Assessment] is very much a little extra thing that we do, it is not our main focus" (SA13).
>
> (Low value and motivation, high cost and negative attitude, leading to little effort in assessment practice)

> "The university does run various workshops and things like that … There are obviously a number of different workshops that are available … some of which I would go to, but the fact remains that the vast majority of convenors and examiners in the clinical years are clinicians … We don't have time to go to workshops, okay, because we run a clinical service. And the vast majority of the teachers are first and foremost clinicians and their teaching commitments are added on top. And so that often precludes one from, you know, being able to go to a lot of workshops and things like that" (SA18).
>
> (Varied impact appraisals, perceived barriers/cost and values, prioritise practitioner identity and work over assessment and HPE)

> "We had to stop doing that [oral assessment] because of the burden of assessment and it was just taking up too much of the examiners time, to be honest, and we didn't find that it added much value" (SA14).
>
> (Low value and motivation, high cost and poor response efficacy, negative attitude, leading to less time and effort for their assessment practice)

"It was hugely labour-intensive and simply not doable for us – [that] is the bottom line. So, we moved, we turned, so that [long answer questions] then became our theory MCQ" (SA18).

(High cost, poor impact appraisal, leading to less time and effort for their assessment practice)

In contrast, participants who experienced a personal interest and reward in practicing assessment, had a high value and motivation to invest and improve the quality of their assessment practice (a positive response appraisal). These factors triggered their seeking of further HPE training in order to develop their assessment literacy and agency which would assist them in these endeavours.

"I love this process. This process began couple years ago … When I was in my residency programme in [names specialty], and during my fellowship, I was quite interested in the way we learn and the way things are taught … As any other attending, while I was taking care of my patients … I interacted with students, with residents, in [names speciality] … The whole renovation of the postgraduate curriculum in surgery needed to be done. [My supervisor] invited me to participate and to get acquainted with the whole process … I've been every so often engaged in educational activities, learning how to do these things … Every time I step into one of those arenas [conferences] I usually take a couple of courses regarding medical education" (MX3).

(Personal interest/high value and motivation, leading to further HPE training)

"I'm very interested in in the education part, so, I try to go to seminars. Every year, there is a congress of education ... and we always send someone to participate, and we take the classes … and different techniques. I'm not an expert, but I try to see some other [HPE] things, to see if something can help [my assessment practise]" (MX7).

(Personal interest/high value and motivation, leading to further HPE training)

"I have a Master's in education … I was sure that just because you are a [clinical] specialist, you are not a teacher. And we tend to feel that we are, especially [names speciality], they

have a big ego, and they know everything, and they know the world. When I started doing my Masters, some people told me, "Why are you doing that? You're wasting your time. No one is paying you more because you have a Master's in education." I finished it two years ago and I think it widened my … point of view and [I] try to get better at doing this [assessment]. I try to make the students get the whole picture of what we want to do … I think it should be like encouraged" (MX10).

(Personal interest/high value and motivation, led to further HPE training and changed practise, despite resistant culture of colleagues/department)

"I think it's my own interest, it's not common, it's not the way we are … It's because it's my field of interest, I was really involved in medical education … I was fascinated by the outcomes, and curriculum design is one of my favourite parts of that. And it was really easy to see that evaluation … [is] valuable. So, I was really fascinated with the idea about this evaluation process, how we could really do it … to see really what is going on with this student is fascinating, how the portfolio could help us to do that better. So, I like evaluation, that's why I was involved in it and I try to keep involved" (MX12).

(Personal interest/high value and motivation, high response impact, led to further HPE training and changed practise, despite resistant culture of HEI/department)

Unfortunately, when one participant, and their team, attempted to completely redesign their assessment practice, implementing new teaching, learning and assessment strategies, they were left feeling demotivated at the lack of the results seen at the end of their course (leading to negative impact and response appraisals). Time and effort put into their assessment practice did not translate into improved student learning outcomes (poor impact/effect), leaving them questioning whether or not they should persist in their new assessment practices.

"What we did do, or what we attempted, was to try and push students in the direction of, you know, of self-, of active-study … to have them do presentations about a topic and we did group work as part of the class mark. They had to present on a topic as a group, so they could learn from one another and then ... so they could research some literature,

and could think, they had to answer questions about it in class … and the class then listened to what they ... you know they had to present in groups … So we tried it, we thought it would … help to push them in that direction, but when we saw the first test *it didn't seem that it had made any difference*" (SA4) – *emphasis* added.

(Varied impact/efficacy and response appraisal, affecting future assessment practice)

A personal factor, specific to **South Africa**, being mindful of South Africa's Apartheid history of racial discrimination, was the acknowledgement of the potential impact of personal biases (related to attitude), on their assessment practice. For example, during oral examinations, students expressed concern of being unfairly judged based on their English-second language ability (accents), raising concerns of hierarchy and victimisation, which was indeed brought to light in the more recent unrest and student protests around #RhodesMustFall and #FeesMustFall, and the call to decolonise HE. This led to an increased reflexivity and sensitivity (awareness and mitigation of potential biases and negative attitudes) of assessors in their assessment practices.

"The other thing is invariably the interactions between … examiners and candidate. Some examiners will lead on more, some will lead on less … and I've little doubt that relates to all sorts of … *cultural and gender biases*" (SA6) – *emphasis* added.

(Potential cultural/ethnic and gender biases)

"It [assessment] has to be fair… It should [be] unprejudiced towards the person whether they are male or female or are in a different language or cultural group than you are … You can't assess another person based on your perceptions … There has to be an … almost uniform standard against which you have to assess people" (SA7).

(Potential gender, cultural/linguistic biases, leading to increased reflexivity and sensitivity in assessment practice)

"One thing that has changed over the last maybe five years … with all the students' … concerns that have come up is being much more aware of students' backgrounds and languages … So, whereas before people might have been penalised because they made

some mistakes with their grammar, now that obviously wouldn't take place … A much greater awareness of where students come from and their background" (SA13).

(Potential bias due to ethnicity and language, leading to increased reflexivity and sensitivity in assessment practice)

"The petty politics that go on, I think, and using the broader vision of what the new South Africa is about … the biggest real problem is in *the minds of the people who are actually in the institution*. That is the biggest problem and in infighting that results in using those sorts of politics and excuses … So, I think though, that although all the other issues are important, it is really the response that academics have towards the politics that has a much bigger influence on our ability to teach and to assess" (SA12).

(Potential racism)

A unique personal factor mentioned in the **Mexican** context was age, or the generational gap, when speaking of using technology in assessment in particular. In general, older individuals were less competent, and more resistant, towards using technology (such as an application on a smart phone to rate a students' clinical performance), than younger individuals. This factor may not have been mentioned by South African participants simply because technology does not yet play a significant role (it is costly, and the required infrastructure, such as reliable internal access, in all clinical settings where students are placed and assessed, remains lacking) in the South African assessment context.

"There are people who help and people who do not. And, I believe that a big part of that is the technology part, because young colleagues use all these apps without any problem, and the older colleagues are teachers who are like, "I don't know how, no, I won't do, I won't do it, I won't do it"" (MX5).

(Age/generational gap, related to electronic literacy, leading to varied assessment practices)

"One of the problems we have is age of the [names specialty], because we have 35-year old [specialists], but [then] we have 65-year old [specialists] that can barely use a

smartphone. And the students get trouble because they are with them 15 days and say, "The doctor didn't evaluate me at all", or the doctors have troubles with the form, or the doctor just gave the student the form to do, "You evaluate yourself because I don't know [how]", [or] "Four, four, four, four." So, I think it's, but going back to previous point, we have a lot of professors in [names speciality], the ages are very wide. So, maybe the young ones are the most friendly users of the app, the old ones are not" (MX10).

(Age/generational gap, related to electronic literacy, leading to varied assessment practices)

To conclude, these quotations illustrate the complexity of assessment practice, for, despite lecturers' conceptions of assessment, numerous additional **personal factors**, along with contextual factors, also impacted on the enactment of assessment. These were: a sense of agency, HPE training and/or formal education, value, motivation, attitude, and age.

7.2. Contextual factors influencing lecturer assessment practice in diverse Southern contexts

"You have a big class and limited time, with a big module to which many different people having to contribute … different inputs … What is practically executable? We don't live in an ideal world. Any system has its faults. And the other question is: what is good enough?" (SA4).

The following **contextual factors** impacted on assessment practice: perceived barriers, such as, competing workloads (clinical, research and educational), and resource constraints (time, human and financial resources), which related to response costs and appraisal, interpersonal, departmental, institutional, national and political factors.

Participants mentioned several perceived barriers, which were related to resource constraints and external costs (competing demands). These included, workload pressures, from a host of other activities, such as clinical, research and educational work. Due to competing workloads, participants often shared a lack of time as a reason for a lack of effort and importance being tied

to their assessment practice. This time pressure was compounded by increasing class sizes (student numbers), often without an increase in manpower or finances. These factors contributed to response costs (high) and appraisal (low, referring to a lack of action, such as quality assessment behaviour, being taken). Factors that further contributed to these response costs and appraisals were identity and associated value. Participants with practitioner identities valued and prioritised their clinical workload and disciplinary knowledge over their assessment responsibilities and HPE training. This was linked to (negative) attitudes towards assessment.

> "It's not the only thing that I need to think about all day … At the end of the day, we are trying to do our best with what we have" (MX11).
> (Perceived barrier/clinical workload, response cost)

> "They have lots of clinical work. It [assessment] isn't high up on the list of priorities" (SA7).
> (Perceived barrier/clinical workload, practitioner identity and work prioritised, low value and motivation)

> "Well, the problem is that I, *in theory*, agree with you *but in practice* the demands on our teaching, especially … the undergraduate load and what we are trying to do, is such that *it is not feasible*, it's just not able to do … there is *so much pressure*, and that pressure is increasing to do all kinds of other things as well; research, getting involved with the Province [government], attend meetings, go to courses that the university thinks are good for us, for the students … and then trying to keep up with professional … reading and interaction … it's really, assessment, quite frankly, if I were to summarise it in one word, a *nuisance*" (SA12) – *emphasis* added.
> (Perceived barriers and resource constraints, high response cost, negative attitude)

> "Limiting factors would be … time, money, budget and of course staff complement, we are a very small division … [and] our people … are [in] joint-posts [clinical and educational] … So they work out at the community health centres and 30%, I think, 30% of their time is for teaching. So, they then have to block off these activities for exams, I don't have a separate component of examiners. So, those are the challenges … We can recruit people

from outside, but then it's got budgetary constraints, because you need to train these people, and of course you have to have workshops, you can't just have people from outside. I would love to have GPs, but they would have to have a postgraduate qualification in family medicine … You have to have training workshops, you have to pay them … Those are the issues … We do have external examiners … [but] budgetary constraints … it becomes expensive" (SA15).

(Perceived barriers and resource constraints, high response cost)

"With the current curriculum you can't reach everyone … And not with those huge classes and lack of manpower … When I teach a class, I don't really teach for assessment purposes. I teach a class with the focus on finishing the curriculum and teaching what I think is the minimum requirements to become a GP" (SA6).

(Perceived barriers/large class size and resource-constraints/lack of manpower, leading to a minimising of their assessment practice)

"There are a lot of students and due to the lack of manpower we sometimes have to do it in a very short manner. And we have to know that we can't cover the whole field … We try to identify what we feel is more important or appears more frequently, then some things are unfortunately neglected … I think when the time factor comes into play… numbers [and] time … it's a big problem … When you have to assess a person you would like to take a whole day and sign him off and then you can say this guy … "He can go." But, you only have half-an-hour, or two hours, to assess the whole year – I think most people would agree with me, that it is actually impossible. That for me is a huge gap" (SA7).

(Perceived barriers/large class size and resource-constraints/lack of manpower, leading to a potentially invalid assessment practice)

"We just don't have that kind of capacity. I run the fifth-year and sixth-year undergraduate programme, I am the sole convenor, so there is a lot that goes into that. I

don’t have the capacity to feedback to each student … I do need to do most of it myself … it takes time” (SA17).

(Resource constraints, high cost, leading to a limited assessment practice/no feedback)

“I told you about the resource constraints … There are just not enough lecturers … An adequate assessment would be adequate patient material and an adequate number of staff. We don’t have that” (SA8).

(Perceived barriers and resource constraints, leading to a limited assessment practice)

“We’re all so busy with so many demands on us, that we can only set a certain amount of time for this … I feel we do not have enough lecturing time … And then we are expected to deliver on research, which from our promotion point of view, in terms of how the staff thinks about it and the most important thing the university is interested in is the research output ... So we’re working at the bare minimum in all these respects” (SA8).

(Perceived barriers/competing workloads, varied response impact and appraisal, leading to less effort for assessment practice)

“The situation isn’t really that good, I have to do both [clinical and educational work] in order to get money to survive. If I am just doing institutional practice, I don’t get enough money to survive. So, it consumes me, [it takes] a lot of time doing both. I also have to do something for the HEI, which [also] consumes me [and takes] a lot of time. When you are doing these three things [clinical, research and educational work], and they ask you to go to a [HPE] course that lasts three hours, usually nobody wants to be there and don’t pay attention at all, answering phone calls from the hospital, from patients, we are not that really into the course … [Clinicians] have to cancel their practice in order to go to those [HPE] programmes. And we don’t get paid by going to these programmes. So, they lose money and they complain about that” (MX2).

(High external costs and low response appraisal, competing work demands and values led to a lack of time and effort for HPE assessment training)

“I had to set the test on my own. When you ask the other guys a question … they're [colleagues] just too busy for this sort of thing. So, it is the same stereotypical questions. There aren’t a group of people that are going to sit down and think, like, “[We] want to do [this]” … A whole bunch of guys, that would be a wonderful, positive thing ... It would help all of us, because I know the problem we have is relevant to all the other places: they have lots of clinical work, [and], it [assessment] isn’t high up on the list of priorities. When the guys say, “Oh gosh… I completely forgot you know”, and you say you want it by tomorrow, then they sit and suck a few things out of their thumb quickly … I wouldn’t say it is well thought out, and they don’t necessarily think of all the different levels of assessment … [To] formulate a clinical problem, where they have to think a bit, takes a lot out of you too … They just don’t have time for it … with the amount and with everyone being very busy and with clinical work and all of those things” (SA7).

(Perceived barriers/competing workloads and values, resource-constraints/lack of manpower, high response cost and low response appraisal, leading to a potentially poor assessment practice and privileging of clinical work)

“It was a *massive workload* for me, because … there isn’t a commitment [from colleagues] … Nobody wanted to dedicate a whole department’s time to one whole day [of assessment], because of their clinical commitments as well. So, I think *commitment to assessment is influenced by availability of manpower*” (SA11) – *emphasis* added.

(Resource constraints/lack of manpower, perceived barriers and high response cost/identity, value, clinical workload and time, leading to a lack of involvement in assessment practice)

“Sometimes, if you don't do it in the moment, then you can forget the details or maybe I sometimes I do it at the end of the day and then I said, "Okay, this guy went to surgery with me, and then this other guy went for the surgery" and it was too much to remember. And it's time consuming … I think it's time consuming … We have big league [name specialty] … they have a lot of work and they're operating all day long and they don't want to [assess] and they say they don't have time … because they are very busy. So, I think it's lost opportunities, because they have a student that is working with them, in two or three

or four surgeries in one day, and they don't get the evaluation they should be getting" (MX10).

(Perceived barriers/time and competing workloads, leading to poor assessment practice)

"They [colleagues] are still not doing their part [in assessment], so it's been hard … There is no way to pressure them, so they [continue to] do poorly … Professors do not grade whatever they have to do … [They are] reluctant, or there's some that have a lot of clinical practice, so they really don't have time" (MX7).

(Perceived barriers/identity, clinical workload and time, leading to resistant colleagues/a lack of involvement in assessment practice)

"[I want] to change a lot of things … [but] it consumes a lot more time than I was expecting it to … that's basically the thing I'm having trouble with … I don't really have that much time right now … I don't know what I am going to do … because I am not going to be able to find that time. I do have to quit something, the institutional practice or the private practice, in order to be able to teach more. I do have to find a decision, but I am not able to find it right now" (MX2).

(Value but high external costs)

"In practice the *demands* on our teaching … and what we are trying to do … it is *not feasible*, it's just not able to [be] do[ne]. There is *so much pressure*, and that pressure is increasing to do *all kinds of other things* as well: research, getting involved with a province [government], attend meetings, go to [educational] courses that the university thinks are good for us, for the students … and then trying to keep up with professional [disciplinary] … reading and interaction … Assessment [is], quite frankly, if I were to summarise it in one word, a *nuisance*" (SA12) – *emphasis* added.

(Negative attitude, value, response costs and appraisal)

High response costs (resource constraints, perceived barriers, and a lack of assessment literacy and agency) and low response appraisal (less time and effort given to the practice of assessment)

was specifically seen in participants prioritising a convenient assessment practice over a thoughtful and critical (HPE evidence-based) model of assessment. For example, participants shared questionable assessment design and scoring choices. This was observed in altering their assessment practices from insightful long answer questions to multiple-choice and short-answer question formats, for the explicit purpose of easy marking, and the assigning of "excellent" scores to every student without properly evaluating them by the appropriate criteria, both of which raised questions of the validity of these assessment practices.

> "MCQs are useful because they *mark automatically* and *short answers are easier to mark* … [they are more] *convenient* for the lecturer" (SA6) – *emphasis* added.
>
> (Response cost and appraisal, leading to prioritising a convenient assessment practice)

> "You want something that's *easy to mark*, quite frankly, that's probably the most [important] ... you sit with 180 exam papers, you know, it's a weekend in your life that's just gone" (SA12). – *emphasis* added.
>
> (Response cost and appraisal, leading to prioritising a convenient assessment practice)

> "The essay type of questions are … *very labour intensive to mark* thoroughly and fairly … with limited numbers it might work, but … they've moved away from it" (SA3) – *emphasis* added.
>
> (Response costs and appraisal, leading to prioritising a convenient assessment practice)

> "I try to move away [from] deeper knowledge questions … I have realised that *it takes much longer to mark*" (SA4) – *emphasis* added.
>
> (Response cost, leading to changing assessment practice from long to short answer questions)

> "I think one of the dilemmas is … simply the fact that student education is one of many things ... one of many balls one tries to keep in the air, and, that this [MCQs] is reasonably

time effective, it's easy to mark … but in the long term it's not necessarily the best way" (SA4).

(Response costs and appraisal, leading to a convenient assessment practice)

"I fall into the mistake of, "Very good: 4, 4, 4, 4", just like giving all the points without really taking the time, not even if there is space to give like special feedback, I rarely use it, I don't know, because of time … Time is probably the main problem" (MX5).

(Resource constraints/time, leading to an unreflexive and potentially invalid assessment practice)

"Sometimes they [colleagues who assess] don't pay attention. That's one of the most important problems … [They] just put, "Excellent, excellent, excellent, excellent" … We have really busy schedules and really busy clinical rotations … we have been getting more students and bigger groups and we are not doing that well" (MX12).

(Resource constraints/competing workloads, perceived barriers/large class sizes, leading to potentially invalid assessment practice)

As previously mentioned, some participants practiced a rules-based assessment, for instance, blindly following the rules and historical practices ("status quo") of their colleagues, departments and HEIs. While, this could be due to a lack of assessment literacy and reflexivity, a naïve acceptance and implementation of said rules, or having a low value assessment, in other instances, it was because of resistance from colleagues, departments and HEIs, and their associated cultures (normative beliefs, values and traditions), that prevented participants from introducing change into their assessment practices.

"I think people grew up with the idea that that is what education is about. It's only later that… [I] really *looked beyond my beliefs* and said, "*That's just how it works*" … I *realized* there is a theory behind this thing … We *learned* about different assessment methods and so there *I started thinking*, and it *didn't necessarily go through to the rest of the department* … If a guy says he's a [clinical] specialist in a field [then] there's nothing you

can teach him, “We’ve done it like that for years, so why would we change it?” … That kind of thing does happen” (SA10) – *emphasis* added.

(HPE training led to increased assessment literacy, however, departmental norms/beliefs led to resistance towards assessment practice change)

“The older doctors just believe, “That didn’t happen when I was young, we just went there, and learned, and that’s the way it’s supposed to be”. The change, I don’t think it’s something that’s helping them so much … Since they are older and most of them were also teachers of me, it’s a little harder to convince them to do these things … That’s the hard part, the battle against those who are doing the same thing that we were doing that the beginning” (MX2).

(Age/generational gap, colleagues norms/beliefs, resistance towards assessment practice change)

“Everything that has a task is seen as a burden, especially with clinicians, that’s true, “Why do you want me to change? If what I’m doing is doing it okay, why do you want me to change?” So there is always resistance … I think the same thing hinders and the same thing [that] helps – clinicians” (MX3).

(Privilege practitioner norms/beliefs/values, resistance towards assessment practice change)

“All of them are multiple choice questions … I believe that comes from our rules from the [names HEI]. We are not allowed to make questions that are open questions … it is not up to me … I asked if I can change the cut-score of the final exam, but, [I am] not going to be able [to], [because] that is something above us … We do [have] governmental [rules that we must follow] … the ones that I told you: if it’s in the objectives you can ask the question, if it’s not on the objectives you can’t ask the questions, the question has to be a clinical case scenario, the question shouldn’t be that long; the student must be able to read it in 40 seconds” (MX2).

(HEI/rules-based practice, leading to constrained assessment practice)

> "We have prescriptions … guidelines for that [assessment] and you try to follow [them] … We are forced to take all of those elements [rules] into consideration, because, if you don't, you get a letter saying you didn't take it into consideration. And you try to follow those guidelines … because it makes it easier for [the HEI], but, also for you, because you don't want to redo everything … If you don't stick to the rules, you'll always have a problem" (SA5).
>
> (Rules-based practice, to avoid additional work/administration)

While colleagues and departments expressed resistance towards changing assessment practice, national culture also contributed to this resistance.

> "The main challenge here is to change the focus of teachers, compared to what they were used to through teaching this curriculum, for many years, but that has [now] changed. So, we need to also change the perspective of teachers to implement this new curriculum … Resistance is common … The main reason would be that we are all from *Latin America*, or *Mexico* ... The *national system [and culture] is very traditional*, you know, every change is scary … They are *scared to experience new things* or technology or to have new technologies they don't know how to use … It's hard, but it means that professors have to be more flexible" (MX6).
>
> (National culture, resistance towards assessment practice change)

In contrast, other participants found their colleagues, departments and institutions to be a source of support for their assessment practice, enabling and empowering them to make changes. For example, one participant encouraged assessment change in their course through their personal investment in assisting others with their assessment practice.

> "I will be introducing a new, programmatic structure of assessment practice … There is sometimes a struggle with resistance, [but], plenty of the faculty believe in me, trust me and they also respect me as a leader … I go into [assessment] business personally … I

personally support people. For example, yesterday I was helping someone with [their] assessment at 10pm" (MX6).

(Positive interpersonal interactions leading to change in assessment practice)

In Mexico, an interesting institutional culture of change was observed (in contrast to the national culture of resistance to change). This culture of change related to the valuing of education by this institution, as opposed to the typical valuing of clinical work and research by medical HEIs. Part of the routine practice of this HEI was a five-year cycle of curriculum change and compulsory educational training (to equip academics in the new system). Resistance from colleagues and departments was found, as described above, however, some found this HEI culture of (forced) change a positive factor in driving change in their assessment practice.

"Our [names HEI] changes the curriculum every five years, it's the rule, to change every five years … There's a [HPE] research team that looks for the new tendencies of other schools, [and we] try to imitate or listen to them, and change every five years. So it's a process of evolution. Our school has their own, like, self-assessment, to see what works, what was good, or what was wrong, and based on those results, they change their [practice] … We are changed here, because we want to be the best" (MX6).

(National culture of change leading to changed assessment practice)

"[Change] in the [disciplinary] field is not easy … clinical rotations are not changing at all, they are more or less the same, for many, many years … In my experience, I would say that this [routine curriculum change] is one of the best things that happens in this institution and in Mexico, because it is not [simply] change for [the sake of] change, I can validate it is because there is [HPE] evidence that what we need to do something [new]. Sometimes we have been the first pioneers in these kind of fields in Mexico, and there is evidence that we need to go over there. So, I think it's good, but this is hard" (MX12).

(National culture of change leading to changed assessment practice)

Political pressures were also mentioned as having an impact on lecturer assessment behaviour, referring to the influence of national politics on HEI assessment practice. For example, in light of the legacy of Apartheid and racial discrimination in **South Africa**, participants expressed feeling pressure from their HEIs to pass students, which was sometimes related to race, as was supported by the government. Public HEIs also relied on the government for funding, therefore, if the government desired to see increased student outputs, particularly students of colour, that added to the pressure felt by assessors. In one instance, a participant shared that they adjusted their assessment results to ease the pressure felt.

> "What is happening at the moment … is this whole pressure to pass because, *"Or else the department doesn't get the money*""(SU12).
> (Political pressure, external costs)

> "We're all affected by politics. Politics is really about decisions to how the money goes … [The] cry "Pass one, pass all" from a *political point of view* is obviously completely misguided and misunderstanding what qualifications are for. But it is also something that the *Department of Education encourages*, because, as in many of the weaker institutions, you've got *financial problems*, if they fail a significant proportion of students they are *financially embarrassed*, so that is a perverse incentive and it makes it very difficult for the previously disadvantaged university … I think that there needs to be a much more subtle, complex, way of *funding and providing money* for … disadvantaged universities … It's pointless for the *Department of Education* to do that and to say well we're putting more *money into student bursaries* but we only pay out the bursaries when the students qualify" (SA12) – *emphasis* added.
> (Political pressure, external costs)

> "The particular person was under immense *pressure* due to the *political climate* to graduate or pass … a *university [and] political pressure*. You can't fail twenty students, only ten are allowed to fail, or no more than five are allowed to fail. [A] student didn't pass and then you get a phone call [from HEI administration] asking, "What you're going

to do now, should we raise his marks, you know?" Then we raise them. When [HEI administration] asks you a question like that … I've never refused" (SA7).

(Political pressure leading to adjusting assessment scores)

In **Mexico**, national pressure differed, yet, also led to a felt pressure by assessors to pass students. Participants interviewed from Mexico assessed at a private HEI (in contrast, all medical HEIs in South Africa are public institutions). In Mexico, participants felt accountable to students, in terms of passing them, because of the high cost of tuition paid by these students to be trained at a private HEI. Whether or not this impacted on these assessors in terms of adjusting student marks, as was seen in South Africa, was not clear.

"Our students, they pay a lot to get into our institution, their dream is to get into residency [pass a national examination] … As the institution, we have to take that in mind and it's our responsibility" (MX4).

(Private HEI/high tuition, accountability to students/pressure to pass)

"I'm paying a lot of money for [names private HEI] for my kid, I would like him to have more of a guarantee for passing the exam … I think we should be taking it into account more heavily because parents are worried, I mean I'm paying and then two years in a row, my kids have failed the exams … You have a six-year career with a very, very good school and you should be able to do the specialties you want, and sometimes they don't" (MX10).

(Private HEI/high fees, accountability to students/pressure to pass)

These national, institutional and financial pressures (high external costs), also lead Mexican participants to proposing the need for a more internationalised assessment practice. They believed that, in order to best prepare their students to pass the residency entrance examination, they needed to look to and adopt assessment practices from America and Europe (the global North).

> "I believe that … we must go to the examination, to the assessment, from … international institutions. That way, I have prepared you … to take the international exams, not my exams … once you've finished school [the residency entrance examination] … It's very hard to get into residency here in this country … [We need to] take exams from abroad … at the international level, and not my school, not my class. So I think that we have to move towards that direction … we have to move the assessment of every single class … [to] international standards, not ours, cause we have to move up … I do believe that the main assessment has to be [at an] international standards … I think we have to assess with an international standards from day one" (MX4).
>
> (National pressure, desire for internationalisation of assessment practice)

To summarise, several **contextual factors** were found to influence lecturer assessment practice, including a number of perceived barriers, such as competing work demands and a related lack of time, resource constraints, due to increasing student numbers, and a lack of manpower and funding, interpersonal interactions, for instance, supportive or resistant colleagues, departmental and Institutional cultures, national factors and political pressures.

To conclude, in agreement with the HBT conceptual framework, multiple personal and contextual factors have been described by lecturers in diverse Southern contexts as having an impact on their assessment behaviour (see **Figure 7.1.** and **Appendix Figures 7.1. & 7.2.**).

Chapter 8: Discussion and Conclusion

This research study sought to explore a key area of silence in the field of assessment, in both HE, generally, and HPE, more specifically – the assessor, from the South, who is responsible for assessment at a course level. In HPE, in particular, the consequences of assessment are critically important for, ultimately, it impacts on patient care and public safety, and, hence, the need to ensure a high quality and valid assessment practice that achieves its desired outcomes and delivers competent health practitioners.

8.1. Study Summary

In order to use assessment to drive learning in a desirable manner, those who are responsible for practicing assessment need to be investigated and understood. To this end, the research questions of this study were, (1) what are lecturers' conceptions of assessment, and, (2) what additional personal and contextual factors impact upon their assessment behaviour (see **Figure 8.1.** for a summary of the results).

8.1.1. Lecturers' Conceptions of Assessment in the South

In line with the Phenomenographic methodology used in this study, a range of conceptions of assessment held by lecturers, in diverse Southern contexts, was found. Analysis proceeded through three stages: the first phase, the pilot study, consisted of data collection and analysis at a single South African site. The confirmatory study consisted of the second and third phases of data collection and analysis, which took place at a second South African site, and a Mexican site, respectively. Each phase of data collection and analysis yielded an Outcome Space describing lecturers' conceptions of assessment.

Conceptions from the first Outcome Space were title **Undirected, Content-focused/Reproduction-directed,** and **Competency and Conceptually-focused/Application-**

directed. These conceptions evolved into **Detached practitioner, Emerging equilibrium,** and **Engaged educator** in the second Outcome Space. The third, and final, Outcome Space described lecturers' conceptions of assessment as **Passive operator, Awakening inquirer, Active owner,** and **Scholar**.

These conceptions of assessment contained various dimensions, including their conceptual understanding of assessment, reason for practicing assessment, scope and scale of assessment practice, assessment literacy, temporal perspective, identity, role and responsibility, degree of reflexivity, accountability and emotional valence.

Each Outcome Space will be discussed below, but, in general, I argue that, while the Outcome Spaces might appear incongruous, they are, in fact, reconcilable and display the iterative process of qualitative research and analysis, spiralling from superficial and descriptive to more complex and theorised. Morse (1994) describes the process of qualitative research ("emerging from the data") as comprehending, synthesizing, theorising and recontextualising.

The third Outcome Space represented comprehension, synthesis and theorisation from all the data collected (recontextualization will take place during model validation and transfer to other settings). Morse (1994) also stated that this process is related to the maturation of the researcher, and, as this study took place over three phases, as time progressed, the researcher matured and the analysis deepened. This led to, for example, recognising related elements across all transcripts, which I have referred to as being "implicitly present" in the previous Outcome Spaces, only to be "fully realised" in the third, and final, Outcome Space.

An evolution, particularly in the complexity and richness of the Outcome Spaces, was observed. The initial Outcome Space, from the pilot study, was the most superficial and simplistic. The second Outcome Space deepened and built on from the first. The final Outcome Space was the richest, as all data was now reviewed and considered together.

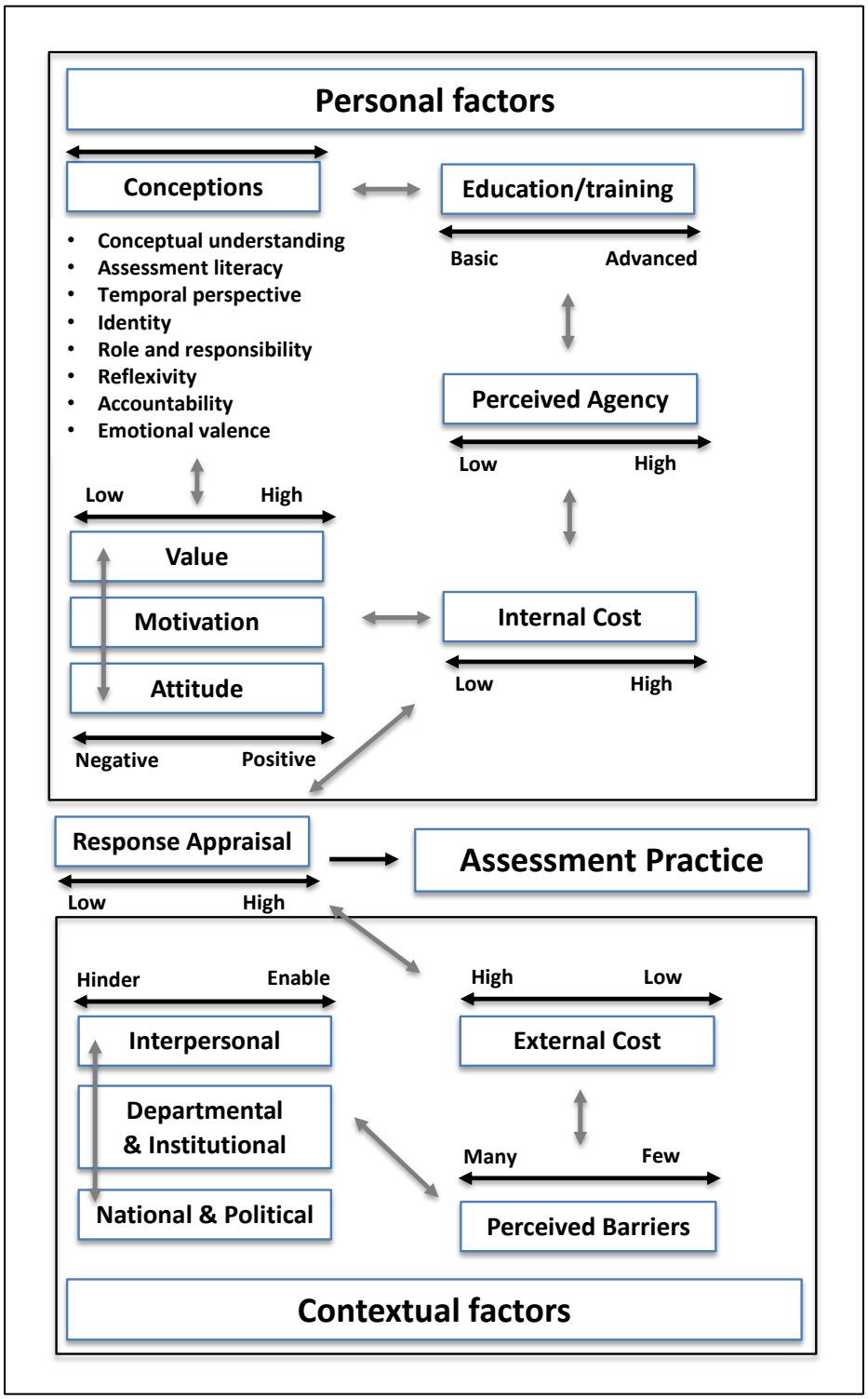


Figure 8.1: Study summary: Lecturer assessment practice in the South. An overview of the findings of this research project. Numerous personal, including conceptions, and contextual factors, interact to impact on lecturer assessment practice.

This evolution was seen in the increase in the number of categories and dimensions. Often, the dimensions related to one another, from one Outcome Space to the next, becoming more nuanced, and refined, with each round of data collection and analysis. The labelling of the conceptions also reflected in increase in complexity.

8.1.1.1. Pilot study: the influence of conception literature

In reflecting on the first Outcome Space, and reviewing literature on conceptions, it became clear that initial analysis was superficial and prescriptive, based on the interview questions and influenced by prior reading on conceptions (see **Section 3.3.1.1.** and **Appendix 3.3.**). The prescriptive nature of the dimensions has previously been mentioned, but the labeling of the conceptions of assessment were informed by literature on conceptions of teaching and learning by Kember (1997) and a study on conceptions of assessment by Postareff et al. (2012).

While conceptions of teaching and learning may not be conflated to conceptions of assessment, in this study, I argue that conceptions of assessment may not be separated from conceptions of teaching and learning, for teaching and learning proceed assessment, and a sophisticated understanding of assessment includes aspects of teaching and learning, and the wider curriculum. For example, the concept of blueprinting in assessment: an assessment practice is planned against learning outcomes, meaning that the assessor needs to take all teaching, learning and assessment activities, the entire curriculum, into account in order to design and implement their assessment (Coderre, Woloschuk, & McLaughlin, 2009; Hamdy, 2006; Wass, van der Vleuten, Shatzer, & Jones, 2001). Indeed, despite asking participants about their conceptions of assessment, they also mentioned conceptions of teaching and learning. While the relationship between conceptions of teaching, learning and assessment was not explored in this study, I do believe there is a case for parallels existing between said conceptions, which other studies on lecturers' conceptions of assessment have claimed (DeLuca et al., 2019; Fletcher, Meyer, Anderson, Johnston, & Rees, 2011; Norton et al., 2019; Wass et al., 2001).

Kember (1997) described a teacher-centred/content-orientated (knowledge transmission) conception of teaching and learning, and the student-centred/learning-orientated conception. Postareff et al. (2012) described a transmission teaching-reproduction conception and practice of assessment or constructivist-transformational conception and practice. Both studies were agreeable with the first Outcome Space, in terms of viewing assessment as a tool to drive superficial, content-reproduction learning, based on what the teacher thought was important, on one end of the range, to a view of assessment as a tool to drive the transformational learning, conceptual change and development of students on the other end.

The description of the "Content-focused/Reproduction-directed" conception of assessment, resonated with what was found in other studies on conceptions (Bolander et al., 2006; Kember, 1997; Murray & Macdonald, 1997; Postareff et al., 2012; Samuelowicz & Bain, 2002; Visser-Wijnveen et al., 2009; Watkins et al., 2005). In these studies, these conceptions put the teacher, and their knowledge, at the centre, and not the student or their learning, but, perhaps, more uniquely to our conception, a development, in terms of their technical knowledge was observed, similar to the "awareness" conception of assessment described by Halinen et al. (2013).

The "Competency and Conceptually-focused/Application-directed" conception of assessment also shared conceptions with Halinen et al. (2013). Halinen et al. (2013) reported a "development" conception, describing a lecturer's progression from being aware, to being able to explicitly articulate sound assessment principles. This idea, of conceptual develop, moving on from a lecturer's personal understandings of assessment, to the view that the purpose of assessment is to support students in their learning, growth and conceptual change, has also been reported in the assessment conception literature (Postareff et al., 2012; Visser-Wijnveen et al., 2009). In these studies lecturers viewed assessment as a tool for developing student competency (Bolander et al., 2006). This lead to a change in focus, shifting from teacher-centredness (content expertise) to student-centredness (prioritising their learning) (Kember, 1997; Samuelowicz & Bain, 2002). Furthermore, this conception of assessment extended towards social and moral

functions, as was found in another study on conceptions of learning and knowledge (Entwistle, 1997b).

For the "Undirected" conception of assessment, Brown (2004) described four possible conceptions of assessment among school teachers, including viewing assessment as "irrelevant", and, while the author did not find any teachers to hold this conception, it could still relate to the Undirected conception of assessment understood assessment to have no to little meaningful function, or, in other words, was irrelevant. Another shared conception could be that of "status quo", reported by Halinen et al. (2013), as these lecturers unquestioningly, and uncritically, followed what had always been done. This was seen in the Undirected lecturer who possessed limited conceptions of assessment and blindly accepted the traditional or historical (status quo) practice.

8.1.1.2. Second phase: identity emerged

During the second phased of the study, identity (discussed in **Section 3.3.1.2.**) seemed to provide the foundation upon which lecturers in HPE based their conceptions of assessment. Indeed, a study suggested that an important factor in influencing the design of assessments was professional identity, and recommended a strategy to improve assessment practice by targeting professional identity development (Bearman et al., 2017).

This concept of identity was seen clearly in the data, for lecturers would directly describe themselves as a "Clinician", and use that descriptor as a core explanation for their assessment understanding, practice and assessment-related decisions[11].

[11] Basing clinical assessment-related decisions on clinical experience and expertise is not necessarily poor assessment practice, for, recent studies have reported the limitations and problems in using objective and standardised tools to measure critical thinking and competency, and, how subjective, but expert judgement, based on many idiosyncratic reasons, can lead to valid decision-making in a clinical assessment setting (Berendonk et al., 2013; Crossley, Groves, Croke, & Brennan, 2019; Govaerts, van de Wiel, Schuwirth, van der Vleuten, & Muijtjens, 2013; Rear, 2018; ten Cate

The differences, and distinctions, between the extreme categories on either side of the spectrum, "Detached practitioner" and "Engaged educator", were clear, but the category between them, "Emerging equilibrium", was less clear, for it shared aspects of both conceptions.

Relatedly, I asked whether, or not, the "Emerging equilibrium" conception of assessment was a distinct category, or, a possible developing and transitory category, whereby, lecturers may pass through this category as they "progress" from "Detached practitioner" towards "Engaged educator"? While Outcome Spaces do not depict the progressive journey or developmental pathway of individuals, and it is beyond the scope of this study, as identity has been described as a dynamic construct that is continually re/formed (Steinert et al., 2019), I could postulate that lecturers could potentially "progress" along and "up" the hierarchy of the Outcome Space (which would be the goal of faculty training), based on time, experience or practice (Cantillon et al., 2019; Ericsson, 1998), interacting with others (Cantillon et al., 2016), training and educational opportunities (Åkerlind, 2003; McLeod, Meagher, Steinert, Schuwirth, & McLeod, 2004; Norton et al., 2019; Steinert et al., 2019; Stenfors-Hayes et al., 2011), or other disorientating event or factors, that were uniquely impactful and meaningful to a particular individual (Kay et al., 2019; Meijer et al., 2016; Mezirow, 1997). However, as this was not explicitly explored, it remains an area of future investigation.

During the development of the "Emerging equilibrium" conception, it was considered whether the clinician and educator identities were indeed equal (balanced), or, if there was a tension (competition) between the two. I asked, "Should this category be split into two categories of "Dominant-practitioner and Recessive-educator" and "Recessive-practitioner and Dominant-educator", or not? Did lecturers associate more strongly towards one of their identities, or, was there a harmony between their dual identities?" I did not find enough evidence to support a

& Regehr, 2018). The authors continued to recommend subjectivity in assessment for it addressed the need for variability, flexibility, adaptability, and resilience, because the contexts in which clinical assessment is practiced is variable (ten Cate & Regehr, 2018).

divide, rather, it appeared as though these lecturers were content to identify as *both* clinicians *and* educators, viewing, and valuing, each identity as similarly important, and necessary, for their assessment practice. This led to the decision to present the middle conception as a single category, consisting of a dualistic identity. There was also support for this in the literature, where medical students and teachers have been described as possessing multiple professional identities, existing along a continuum of "compartmentalised" to "merged" (Monrouxe, 2010; Steinert et al., 2019).

While identity did not emerge from the pilot study, it was not asked of participants, it was found in the third phase of the study, and included in the final Outcome Space as a dimension.

8.1.1.3. Final Outcome space: Assessment in the South

I have already described how the third, and final, Outcome Space reflects the deepening analysis and maturation of the researcher over time, representing all the data collected across three diverse Southern sites, in its nuanced and complex ("fully realised") forms. How this final Outcome Space relates to the current literature on conceptions needs to be explored.

I have already discussed how the conceptions and dimensions of first Outcome Space compared to broader literature on conceptions (teaching, learning, knowledge and assessment), and, not surprisingly, similarities existed between the final Outcome Space and conception literature (see **Table 8.1.** for an overview). However, in general, the conceptions and dimensions described in the final Outcome Space far surpass current literature on teachers' conceptions of assessment, in terms of depth and detail. The final Outcome Space from this study represented a diverse range of conceptions, with a greater number of components, that contributed a new level of richness to the field.

Table 8.1.: Final Outcome Space commentary: Similarities found between the conceptions of assessment described in this study and the existing literature on conceptions of assessment teaching and learning, expertise and academics educational scholarship.

	Passive operator	**Awakening inquirer**	**Active owner**	**Scholar**	**Reference**
Conceptions of assessment	Assessment as "irrelevant"				Brown (2004)
	Assessment for "school accountability"		Assessment for "student accountability" and "improvement" of student learning and teaching		
	"Status quo"	"Awareness"	"Development"		Halinen et al. (2013)
	Assessment as "transmission teaching" (content-reproduction)		Assessment as "constructivism" (student-development)		Postareff et al. (2012)
	Assessment as "knowledge production" (teacher-centred)		Assessment as "knowledge construction/transformation" (learning-centred)		Samuelowicz & Bain (2002)
	Psychometric		Socio-constructivist		de Jonge et al. (2017)
	Passive operator	**Awakening inquirer**	**Active owner**	**Scholar**	**Reference**
Conceptions of teaching and learning	Teacher-centred/content-orientated (knowledge transmission)		Student-centred/learning-orientated (learning facilitation)		Kember (1997)
	Teacher-focused/ information-transmission		Student-focused/ conceptual change = deeper learning approach		Trigwell, Prosser, & Waterhouse (1999)
	Teaching and learning as reproducing information ("academic")		Teaching and learning as seeking meaning ("social")		Entwistle & Peterson (2004)

	Acquisition of knowledge/ teacher-focused		Conceptual development/ critical thinking/ teacher-learner interaction	Conceptual change teacher-learner interaction	Visser-Wijnveen et al. (2009)
	Passive operator	**Awakening inquirer**	**Active owner**	**Scholar**	**Reference**
Expertise and educational scholarship literature	"Novice" from five stage model of skills acquisition and expertise	"Advanced beginner"	"Proficient" to "competent"	"Expert"	Dreyfus (2004)
		"Every day reflective teacher" to "reflective teacher"	"Reflective teacher" to "scholarly teacher"	"Scholarly teacher" to "teaching scholar"	van Schalkwyk et al. (2013)

In comparing the current body of assessment conceptions literature to the final Outcome Space, similarities were observed. These included the view of assessment as psychometric and constructivist (de Jonge et al., 2017), relating to the summative and formative purposes of assessment, and the concepts of teacher-centred/content-reproduction and learner-centred\development-focused assessment (Postareff et al., 2012; Samuelowicz & Bain, 2002), which related to both the purpose of assessment, and the assessor's role in the final Outcome Space. The element of "status quo" (Brown, 2004) was observed in the final Outcome Space's dimension of accountability (towards the HEI or discipline). The aspects of "awareness" and "development" (Brown, 2004) were present in the hierarchical nature of the conceptions in the final Outcome Space, and in the dimensions of assessment literacy and reflexivity. The dimensions of assessment purpose, assessment literacy, identity, role and responsibility, reflexivity and accountability also relate to the "irrelevant" conception of assessment by Brown

(2004). Overlap between assessment conception literature and the final Outcome Space provides support for the conceptions described, for the conceptions of assessment described in this study may now build upon, and expand, currently available evidence and theory on lecturers' conceptions of assessment.

However, many of the dimensions of the conceptions described in the final Outcome Space were deeper and richer, or unique, to the conceptions of assessment (and teaching and learning), previously described. For instance, the dimension of purpose, in this study, extended to include additional reasons for practicing assessment (moral – accountability to the patient) and unique conceptual understandings of assessment itself (its scope and scale), which tied into the novel dimension of temporal perspective. A "teacher-centred" role (along with a "content-reproduction" assessment purpose) might have been described in assessment conception literature previously (Postareff et al., 2012; Samuelowicz & Bain, 2002), however, in the final Outcome Space, its expansion to include an administrative and operational nuance was new, as was its relationship to a practitioner identity and degree of responsibility (ownership). Similarly, "accountability" to school and student learning had previously been mentioned in assessment conception literature (Brown, 2004), yet, in this study, the dimension of accountability deepened to include aspects of identity, reflexivity and assessment literacy. To illustrate this, practitioners, with limited assessment literacy and reflexivity, used assessment as a disciplinary gatekeeping tool because of a felt accountability towards their profession.

In considering the labeling of the range of conceptions represented in the final Outcome Space, literature on expertise and academics' educational scholarship was consulted (see **Table 8.1.**), because, both areas related to the idea of a developmental journey, which was similar to the hierarchical nature of the conceptions of assessment in the final Outcome Space. I also proposed the idea, in light of the dimension of identity, of the possibility of lecturers transitioning through the Outcome Space, from basic to advanced (as would be the goal of faculty assessment training).

The final Outcome Space displayed the hierarchical nature of a Phenomenographic Outcome Space, for, the first conception of assessment began with the most limited conception of assessment (Passive operator), which then progressed to a more developed conception (Awakening inquirer), followed by a more advanced conception (Active owner), to the final conception (Scholar) at the "top". These conceptions resonated with the Dreyfus (2004) five stage model of skills acquisition and expertise research (Novice, Advanced Beginner, Proficient, Competent and Expert), and van Schalkwyk et al. (2013) journey of growth of academics into educational scholarship (the Everyday reflective teacher, Reflective teacher, Scholarly teacher and Teaching scholar).

The first study described a five-stage model of the acquisition of expertise, defining expertise as an automatic and unconscious response (Dreyfus, 2004). A "novice" had zero experience with a situation, and poor discretionary judgement, developing competency over time, until they had become an "expert", who was able to instantaneously understand multiple abstract components and perspectives, and identify the core solution, or action to be taken, without wasteful consideration of a multitude of situations (Dreyfus, 2004). The second study also described a transition of academics, in terms of their journeys of professional learning and educational scholarship (van Schalkwyk et al., 2013). This transition was seen in teachers who moved from reflection, to teachers who then begun to question, critique and debate practice, going so far as to then performing, and contributing to, educational scholarship, for example, through conducting HPE research (van Schalkwyk et al., 2013).

This work was conducted in Southern settings, which differed, in some significant ways, from those under which HPE is practiced, and researched, in the North. These differences included resourcing, colonial histories and cultural diversity. How these appeared to influence conceptions was seen in the dimensions of emotional valence, such as having a negative attitude towards assessment, because of a lack of resources needed to practice assessment, and, in what I am calling the colonial way of thinking, as displayed by participants in Mexico, through their desire to "internationalise" their assessment practice.

While I may be calling this valuing of internationalisation as "colonial", Southern theorists refer to this idea (that research and theory produced in, and reported from, the West, or global North, is superior, they are the custodians of theory and knowledge) as "Western cultural imperialism" (Gosselin et al., 2016), "Adherence to the dominant discourse" (Montgomery, 2019), and, "Perpetual intellectual servitude" (Morrell, 2016). Indeed, other studies have shown this to be true, where Northern theory and practice were valued more highly, and reproduced in non-Northern contexts, than indigenous knowledges and practices (Beigel, Gallardo, & Bekerman, 2018; Berry & Taylor, 2014; Connell et al., 2017; Connell et al., 2018a, 2018b; Gosselin et al., 2016; Greysen et al., 2011; Montgomery, 2019). This is not to say that assessment theory and practice developed in the global North holds no value for assessors in the global South, but, that caution, and critical reflexivity, need to be practiced before implementing "international" assessments.

While the findings of this study may not radically differ to the current knowledge on lecturers' conceptions of assessment, or provide opposition to the dominant discourse, it is still important to note that this is the *first* study in which lecturers in the South were interviewed, by a Southern researcher, about their experiences, conceptions and practice of assessment. Furthermore, diversity and democracy, in the form of the amplification of peripheral voices and perspectives, has now been added to the theory on conceptions of assessment, which were absent before, which has contributed to the creation of a more "true" model of knowledge (Connell, 2014). Moreover, researchers and practitioners in the South may now consult this research, as opposed to having to rely solely on the intellectual dominance of, and dependency on, the North. Now, no translation from the centre to the periphery needs to take place, because the theory produced in this study, was done in the South, for the South, by the South (Comaroff & Comaroff, 2012; Connell, 2014; Morrell, 2016). Furthermore, if Southern Theory aims to re-centre the South as valid knowledge producers, and not just consumers, or sources, of data, and to amplify alternative and indigenous knowledges, then, I argue that this research contributes to Southern

Theory because knowledge production has been redistributed (Morrell, 2016; Shay & Peseta, 2016).

8.1.1.4. Lecturer Assessment Practice in the South and Factors influencing Assessment Practice

Lecturer assessment practice was the focus on this study. In the global South, the labour and burden of responsibility of assessment typically falls on the shoulders of an individual (course convenor). This means that the course convenor is responsible for assessment from its conceptualisation, design, appraisal, practice, decision-making and evaluation. In contrast, the emerging practice of programmatic and competency-based assessment, in the resource-rich global North, is for competency committees and assessment teams to collective share responsibility. The implications of misguided or poor HPE assessment practice in the global South may be more likely, as no competency committee stands between the individual assessor and their assessment practice, which could then lead to severe consequences for the student and patient. For this reason, the individual lecturer, the course convenor who leads assessment in their block, was the unit of analysis in this study, as the course convenor is an important unit of intervention in any attempts to change assessment practice.

As an eventual outworking of this research is to design interventions to change assessment practice, through targeting the assessor, so a conceptual framework that explains, and may be used to change behaviour, was required. HBT served this purpose, as in seen in the model of lecturer assessment behaviour that was developed (see **Figure 8.1.**). However, as an eclectic and general model of HBT was used in this study, as opposed to selecting and adapting a specific HBT, with mechanisms of action already present, this resulted in a more generic description of lecturer assessment behaviour. Yet, I argue, that a generic model of lecturer assessment practice may lead to greater access and utilisation in other contexts. In fact, I propose that, using a general model, allowed for a more credible qualitative investigation, as the research questions, data

collection and analysis, were not limited to the set factors and mechanisms present in a specific HBT. Rather, the general model allowed space for deep and wide exploration.

In concurrence with the literature reviewed (see **Sections 3.3.1.** and **3.3.2.**), several personal and contextual factors impacting on lecturer assessment behaviour, across diverse Southern contexts, were identified. These included perceived agency, education and/or training, value, motivation, attitude, response costs and appraisals, perceived barriers and resource constraints, interpersonal, departmental, institutional, national and political factors. In reality, it was difficult to isolate singular personal and contextual factors, for, they interacted with one another to drive lecturer assessment behaviour.

In this study, there was evidence that the personal factors of perceived agency and education and/or training (assessment literacy) were related. Participants who had no HPE training often expressed feeling unconfident in their assessment practice, and disempowered to make changes themselves. Whereas, participants with HPE training possessed the abilities to enact quality, evidence-based assessment practices. Moreover, individuals who lacked education and agency turned to their colleagues, with HPE education, for assistance with their assessment practice. This is line with what others have reported: a need for HPE training of lecturers, which can effectively lead to changed assessment practice (Goos & Hughes, 2010; Huwendiek et al., 2010; Karthikeyan et al., 2019; Norton et al., 2019; Norton et al., 2013; Popham, 2009).

The other personal factors, such as value and motivation, which were linked to the conceptual dimensions of identity, role and responsibility, and accountability, were not explicitly found in any of the literature reviewed. However, the personal factor of attitude, which I related to emotional valence, could be related to emotion, for, participants expressed positive and negative feelings towards assessment (linked to value and motivation), which contributed to their response appraisal and action taken, similar to what was found in another study (Myyry et al., 2019).

Response costs impacted on participant response appraisals, which in turn effected assessment practice. Response costs were often linked to perceived barriers, such as resource constraints. Participants, from this study, described resource constraints including competing work demands (clinical, educational/assessment and research duties), which was compounded by large class sizes, a lack of manpower and finances, and, a related lack of time. These high costs led participants to place little time and effort into their assessment practice, as was seen in changing long-answer questions to shorter MCQs for the express purpose of requiring less time to mark (Gilles et al., 2011; Norton et al., 2013; Price et al., 2011).

Just as the literature reported, so colleagues, departments and institutions, their cultures (normative beliefs, values and traditions), impacted on assessment both positively, supporting change in assessment, and negatively, resisting said change (Bearman, Dawson, Boud, et al., 2016; Halinen et al., 2013; Harrison et al., 2017; Jessop & Tomas, 2016; Johnson, Scholes, & Whittington, 2005; Kogan et al., 2017; Norton et al., 2013; Segers & Tillema, 2011). Interestingly, Mexican participants mentioned that a resistance to change was part of the Mexican culture, which concurred with Hofstede's model of national culture, which gave Mexico a high uncertainty avoidance score (82/100), compared to South Africa's neutral rating in the same dimension (49/100) (see **Appendix 4.1.**).

Unique to the contexts of the Southern sites sampled in this study were, national culture, political pressures and the type of HEI, which related to one another. The difference between the public (South African) and private (Mexican) HEIs became apparent through the varied pressures felt, by participants, in needing to pass students. In South Africa, in light of its political history of racial discrimination, accessibility to HE, and throughput, of students of colour, in particular, remains unequal, compared to privileged white students. The government has placed a pressure on HEIs to improve graduation rates of students of colours, which could be tied to the provision of funding from the state. This led participants to feel pressured in passing students, or adjusting marks as dictated to them by their HEI. In Mexico, however, at private medical HEIs, participants felt a pressure to pass students because of the high tuition rates, associated with the expectation

for a quality education and student success (such as, entry into the postgraduate residency programme).

In Mexico, the desire for internationalisation, which I have already described as a "colonial" way of thinking, is also related to the national context: few residency places exist for the large number of medical graduates, which compounds the pressure felt to internationalise (raise) the standards of their assessment practice. The rationale for this belief, as given by the participants, is that the quality of international assessments are higher, and, if adopted in their HEI, their students will be more likely to pass the entrance examination and gain entry into the residency programme. As compulsory clinical training in the public health care system exists in South Africa for medical graduates (two years of internship, followed by a year of community service), South African participants did not express a desire for internationalisation, but rather prioritised a Primary Health Care approach to their assessment practice.

To summarise, the findings from this study confirm, and are supported, by what has been reported in literature, and, rich and deep conceptions, and their dimensions, as well as, unique personal and contextual factors, along with a novel model of lecturer assessment practice in diverse Southern settings, has furthered research in the field of HPE, assessment and Southern Theory.

8.2. Limitations of this study

The focus of this study limited what was explored, interpreted and found. As data collection and analysis took place over three phases, and the desire was inductive analysis, upon reflection, the methodological strategy employed in this study was inherently deductive.

It was true that each data set was initially considered inductively, and in isolation, before deductively comparing the interpretations to other data sets and literature. However, while reflexivity was practiced, and a conscious effort made to temporarily put aside prior

interpretations, they remained sensitizing concepts. In other words, as I was so intensely immersed in the analytical process, it was difficult not to seek to extend and deepen interpretations that had begun to form from one data set to the next. It was not confirmation bias per say, rather the highlighting and following of specific directions. Despite the practice of reflexivity, this could have led to a narrower analysis and the potential missing out of additional dimensions of conceptions.

Since data collection and analysis took place over three phases, as new dimensions and conceptions developed, so, new questions were iteratively asked in subsequent interviews. As identity did not emerge in the pilot study, indeed, it was not asked, yet was mentioned in the first interview of the second phase of data collection and analysis, it was then asked in the following interviews. While the transcripts from the pilot study were reviewed after the second phase of the study was complete, to search for any potential mention of identity, ideally, I should have returned to interview the lecturers a second time to confirm, or deny, the dimension of identity.

In terms of utility, in accordance with the properties of qualitative research, the findings from this study may not be generalisable to other contexts, but they may be transferable. Transferability and utility to other contexts from this study may take place as the results from this study are based on thick, rich, contextualised and coherent descriptions, which have been confirmed and found to be credible in multiple diverse contexts, and a clear, detailed and transparent description of the research process, which may be repeated in other contexts, has been provided in this thesis.

While the eventual outworking of this study is to use the description of conceptions and factors identified, as potentially important influences in shaping lecturer assessment, that may then be specifically and strategically targeted, through the design and implementation of faculty training programmes, it is beyond the scope of this study. This research project also did not set out to determine the strength of possible causal relationships, such as to what degree conceptions may

impact on lecturer assessment practice, or if other personal and contextual factors are of particular importance. Behaviour is complex and changing it is difficult. However, I still propose that, now knowing the range of ways lecturers experience and understand assessment, and what factors afford or constrain their assessment practice, are critical in seeking to change lecturer assessment behaviour. Whether or not an altered lecturer assessment practice leads to powerful learning environments and desirable student learning outcomes, remains to be determined.

Specific issues around the use of the conceptual frameworks, research paradigm, methodology, qualitative research interviews, reflexivity and rigour in this study have been dealt with previously (see **Chapter 4**).

8.3. Utility and future directions

The research undertaken in this study, in exploring lecturer assessment practice, in three diverse, Southern contexts, has led to the development of an Outcome Space that described four conceptions of assessment, consisting of eight dimensions, and, the creation of a model explaining lecturer assessment behaviour (see the summary in **Figure 8.1.**). These findings, while needing to undergo further validation, may serve as a resource for faculty assessment training.

This study, and literature, suggest that both conceptual change, and the consideration of other influencing factors, are needed in order to meaningfully and sustainably change assessment practice (Bearman et al., 2016; Brown, 2004; Carless, 2007; Deneen & Boud, 2013; Harrison et al., 2017; Ho, 2000; Ho et al., 2001; Kane et al., 2002; Kember, 1997; Murray & Macdonald, 1997; Offerdahl & Tomanek, 2011; Pedrosa-de-Jesus & da Silva Lopes, 2011; Watkins et al., 2005). These conceptions, and model of lecturer assessment practice, may then be used to guide the design of interventions and training programmes. For example, through developing a questionnaire inventory to determine where lecturers are conceptually positioned in the Outcome Space, before specific dimensions of said conceptions are then targeted. Similarly, once the model has been adapted to its context, the personal and contextual factors in play may be

targeted during interventions. Whether or not this would contribute to the creation of powerful learning environments, would then remain to be determined.

However, before these findings may be utilised, further validation of these conceptions, and model of assessment practice, need to take place. As the nature of qualitative research is that it is non-generalizable, yet transferable, and useful in other contexts, in order to further strengthen and confirm the findings described in this thesis, the aim is for additional data collection to take place.

Data has already been collected from two distinct samples in South Africa and one in Mexico. The desire is to continue validation of the conceptions and factors identified in this thesis through further data collection and analysis in other diverse Southern sites, such as Indonesia, India, Egypt, Namibia, Zambia and Chile where collaborators are based, and, thus, provide an even richer, robust and rigorous representation from the global South. This research could then be extended to other disciplines outside of HPE, and, to Northern contexts.

The priority would be to first validate the findings of this research in the global South, and within the context of HPE, due to the related critical consequences of poor assessment practice in HPE, and a lack of quality-control assessment committees, before extending to other disciplines and the global North. This is important because of the paucity, and need, for Southern theory (and practice), especially compared to the abundance of available resources, researchers, theory and practice in, and from, the global North. This could also sensitize researchers, both in the global South, and worldwide, to the concept of Southern theory, and encourage critical, socially just thinking globally, as well as inspire researchers from the global South to pursue, and further, the open field of Southern research.

It would also be interesting to explore whether, or not, the proposed Outcome Space represents a progressive journey, a distinct trajectory, for lecturers (Holmboe et al., 2011), for instance, moving from Passive operator to Active owner, and what precipitates that movement.

In terms of validation, a questionnaire inventory, based on the conceptions described in this study, needs to be developed for dissemination. The strength of various relationships and mechanisms of action, as described in the model of lecturer assessment behaviour, also needs to be determined.

8.4. Conclusion

This thesis sought to explore lecturer assessment practice in diverse Southern contexts, specifically investigating what factors influenced their assessment behaviour, and, as a sub-study, what conceptions lecturers held concerning assessment, for, both conceptions of, and factors impacting on the assessor, needed to be identified and understood before any intervention to enhance assessment may be developed.

In terms of the utility of this research, the proposed eventual outworking is to use the findings generated in this study to design an evidence-based and theory-informed intervention to change lecturer assessment practice, improving student learning and health outcomes. The results from this study may be able to help faculty developers to better understand lecturers and their assessment practice in context. The conceptions, personal and contextual influencing factors described provide specific and strategic points that may be target in faculty training.

A hope is that, through using the conceptions described, and the model of lecturer assessment practice developed here, a more impactful, effective and sustainable approach to changing to lecturer assessment practice, improving student learning outcomes and, ultimately, bettering patient care, may be observed.

Appendix

Chapter 3: Literature Review

Table 3.1.: An overview of major assessment instruments, tools or methods used in medical programmes (S. Downing & Yudkowsky, 2009; Schuwirth & van der Vleuten, 2010).

Assessment instrument	Aim	Format
Written examinations	Recall and apply content knowledge	Varied: Multiple-choice questions (MCQs), Short-answer questions (SAQs), essay-questions
Oral examinations	Recall and apply content knowledge, test reasoning and transfer skills	Verbal performance measured by a panel of examiners
Observed Structured Practical/Clinical Examinations (OSPEs or OSCEs) and Simulated or Standardised Patients (SPs)	Test practical or clinical skills	• Rotations between stations of structured cases (clinical situations, tasks or patients) • Performance measured against a checklist or rating scale
Workplace-Based Assessment (WBA)	• Test authentic, real-world experiences and patient encounters • Test professionalism, communication, clinical and teamwork skills	Competence assessed by direct observation in the workplace
Portfolios	Educational instrument	Collection and collation of assessment tasks (reflection, feedback, self-assessment, learning plans and goals)

3.1. Quality control criteria for designing a technically sound assessment.

The criteria for technically sound assessment include reliability, validity, educational impact, practicality, feasibility or cost effectiveness and acceptability (Norcini et al., 2011; Schuwirth & van der Vleuten, 2011a; Shumway & Harden, 2003; Tavakol & Dennick, 2017). It is important for these criteria to be defined as they serve as a form of assessment quality control, standards to hold assessors accountable and used to improve and align all curriculum components, such as teaching, learning and assessment practices, that authentically mimic the workplace and are not divorced from the local context (referred to as blueprinting), as well as ensure rigourous, justifiable and valid judgements, so as to ultimately protect patients and the public (Hamdy, 2006; Norcini et al., 2011; Tavakol & Dennick, 2017; Villarroel, Bloxham, Bruna, Bruna , & Herrera-Seda, 2017; Wass et al., 2001).

Reliability refers to the reproducibility or consistent distinguishing ability of an assessment: similar outcomes or measurements are obtained from the same (test-retest) or very similar (parallel) assessments taken by different candidates on different occasions and may be assured through a sufficiently large sampling of questions; increased scope decreases subjectivity (S. Downing, 2004; Schuwirth & van der Vleuten, 2011a; Shumway & Harden, 2003).

Validity describes how much an assessment is actually and accurately measuring what it purports to measure or, “The degree to which empirical evidence and theoretical rationale support the adequacy and appropriateness of inferences and actions based on test scores or other modes of assessment” (Schuwirth & van der Vleuten, 2011a). There are many different types of validity including, for example, content validity or the extent to which an assessment represents the whole domain of knowledge covered. Validity may be increased through the extent of the assessment in terms of adequate and appropriate sampling in which all domains and levels (from Miller’s pyramid) are represented and tested, as well as selection of the most appropriate test format (S. Downing, 2003; Schuwirth & van der Vleuten, 2011a; Wass et al., 2001). Educational impact, sometimes referred to as consequential validity, speaks to how an assessment impacts

on student learning, such as what, how and why students learn; this may occur through the content, format and timing of an assessment (Schuwirth & van der Vleuten, 2011a; Shumway & Harden, 2003). Another author expands the criteria of educational impact to additionally including "catalytic effect" or the supportive feedback that seeks to drive learning in a desirable way (Norcini et al., 2011).

Practicality, feasibility or cost effectiveness are concerned with the practice of assessment, what will be assessed and how, and thus relies on available resources, infrastructure, expertise and so on (Schuwirth & van der Vleuten, 2011a). Acceptability raises the issue of compromise between various stakeholders, innovation, educational theory and research, as well as institutional beliefs and traditions (Schuwirth & van der Vleuten, 2011a).

A further technical consideration with assessment design is standard setting, or the cut-off value, that determines the consequences, such as "50%" being the cut-score for passing or failing an assessment. The issue lies around determining what is the true value of this standard: is the set standard relative (norm-referenced) or absolute (criterion-referenced)? With relative standard setting (norm-referencing) students serve as their own reference or, in other words, students' performances are described or ranked relative to another student; technically the cut-score is a group of students' scores minus one standard deviation (McKinley & Norcini, 2014; Schuwirth & van der Vleuten, 2011a; Wass et al., 2001). A problem with this method is that it assumes the stability and competence of the group of students, the abilities of the students are not taken into account (Wass et al., 2001). Absolute standard setting (criterion-referencing), on the other hand, determines mastery and uses experts as their reference, in that the cut-off value is a value of minimum competency that experts have determined considering how a borderline student would answer a question correctly or the minimum knowledge, attitudes and behaviours a student must possess (Schuwirth & van der Vleuten, 2011a; Wass et al., 2001). Relative and absolute standard setting are also denoted as norm- (relative) or criterion- (absolute) referenced measurements, whether test scores are interpreted relative to an established norm group (what students know; norm-referenced) or relative to actual content (what students should know:

domain-, objective-, content- and construct-referenced) (S. Downing & Yudkowsky, 2009; McKinley & Norcini, 2014; Tavakol & Dennick, 2017). However it has been argued that standard setting is an arbitrary assumption made about a required level and as long as any standard is rigourous, fair, explicable, defensible, stable or consistent and aligned with curriculum outcomes, then it is acceptable (McKinley & Norcini, 2014; Schuwirth & van der Vleuten, 2011a; Shumway & Harden, 2003).

3.2. Literature on conceptions and practice.

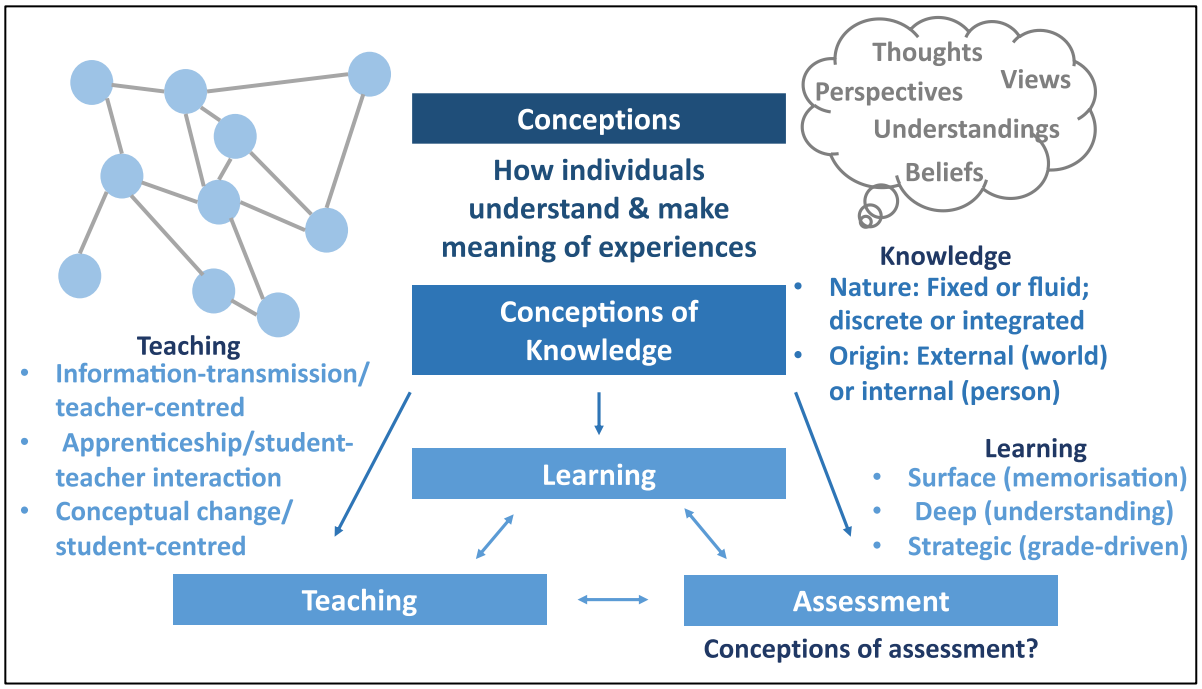


Figure 3.1.: Conceptions: an overview. Conceptions have been defined as varied descriptions of lived experiences or how individuals experience, understand and make meaning of said experiences or phenomena (Åkerlind, 2005; Ashworth & Lucas, 1998; Cibangu & Hepworth, 2016; Dortins, 2002; Entwistle, 1997a; Entwistle & Peterson, 2004; Limberg, 2000; Marton, 1981; Pang & Ki, 2016; Rovio-Johansson & Ingerman, 2016; Sjostrom & Dahlgren, 2002; Tight, 2016). In this study lecturers' conceptions of assessment will be investigated, yet, these conceptions are also related to conceptions of teaching and learning, all of which all are related to conceptions of knowledge and so on. This is in line with Brown's (2004) statement that conceptions are multifaceted and interconnected, like nodes in a network.

Conceptions are argued to be a personal factor of significance in this study because there is evidence that conceptions are important in influencing behaviour and behaviour change (Arcila, 2018; Box et al., 2015; Meijer et al., 2016). In investigating lecturers' conceptions of assessment, teachers' conceptions of teaching and learning need to be reviewed, but, first, the underlying conceptions of knowledge, or epistemological beliefs, also need to be known as they influence conceptions of teaching, learning and assessment. Indeed one study found statistically significant correlations between teachers' conceptions of knowledge and their conceptions of research and teaching (see **Figure 3.1.**) (Visser-Wijnveen et al., 2009).

When it comes to conceptions of knowledge, there are two major elements: the *nature of knowledge* and the *nature of knowing* (Visser-Wijnveen et al., 2009). Regarding the nature of knowledge, questions include asking whether knowledge is fixed or fluid, discrete or integrated; and in terms of the nature of knowing, this includes its source or origin, either the external world or internal world (Visser-Wijnveen et al., 2009). In the Visser-Wijnveen et al. (2009) study the authors described a number of conceptions of knowledge, research and teaching: the dimensions of the conceptions of knowledge ranged from viewing knowledge in terms of information about the external world to being internally constructed, from isolated or discrete facts, to an ever-growing or developing body of information. Conceptions of research included research as disclosing, searching for, interpreting or creating patterns about knowledge either from the external or internal worlds, as well as the role of the researcher as either present or absent (Visser-Wijnveen et al., 2009). Conceptions of knowledge and research focused either on the process or the person, the external or internal worlds.

The literature on teaching and learning conceptions is relatively well-known in education circles, specifically the review by Kember (1997), where conceptions of teaching are described along the continuum of information-transmission/teacher-centred to apprenticeship/student-teacher interaction to conceptual change/student-centred. At one end of the spectrum teachers play the lead role in passing structured information on to their students, moving towards facilitating understanding and intellectual development at the other end, where the student is the

responsible party for learning (Kember, 1997). Indeed, another study described teaching conceptions along a similar spectrum: teacher-focused or teacher/student interaction-focused; and varied in their teaching aims: acquisition of either knowledge, knowledge or skills, or attitudes/conceptual development, or conceptual change (Visser-Wijnveen et al., 2009). Kálmán et al. (2019) state the relationship between teachers' conceptions of teaching and their practice as such, "Teachers' conceptions of teaching have a strong impact on (their) approaches and practices, and, because of this, teachers do not adopt approaches to teaching that reach beyond the sophistication of their conceptions."

Bolander et al. (2006) describes two perspectives teachers hold regarding expertise: an acquisition metaphor and a participation metaphor. The acquisition metaphor is cognitively focused and emphasises information transfer, whereas the participation metaphor is socially focused and emphasises the importance of communication, community of practice and context (Bolander et al., 2006). When asking teachers to reflect on their teaching goals in light of core curricula, the authors found three orientations: content-oriented, competency-orientated and attitude-orientated (Bolander et al., 2006). The study found that teachers with a content-oriented approach to teaching, or implementing the core curriculum, operated within the acquisition metaphor, believing that development of expertise in students comes via content transfer (Bolander et al., 2006) – a similar idea to the teacher-centred, information-transmission conception of teaching and learning (Trigwell et al., 1999). Teachers with a competency-oriented approach focused on developing a handful of isolated competencies in their students by moving away from knowledge to emphasising critical thinking (Bolander et al., 2006). In contrast to the other orientations, teachers with an attitude-orientated approach were, "Concerned with development of the person, their identity, motivation, interest and attitude," and operated in a more participation metaphor manner, focusing on the process over the content, taking the community of practice and context into consideration (Bolander et al., 2006).

Chapter 4: Methodology

4.1. Hofstede's and GLOBE's Models of National and Organisational Culture highlight the cultural differences between countries and shows the need for diverse sampling.

Close to fifty years ago Hofstede conducted a large-scale international investigation on organisations, collecting data by questionnaire from multi-national IBM employees, at all employment levels, in forty-different countries (1967-1973), and concluded that organisations were cultural-bounded, eventually leading to the design of his six dimensions of culture (Hofstede, 2001; Hofstede, Neuijen, Ohayv, & Sanders, 1990). Hofstede's six dimensions of culture are: power distance, individualism, masculinity, uncertainty avoidance, long-term orientation and indulgence; and the countries investigated were given scores for each of the dimensions and a profile created (see **Figure 4.1.** & **Table 4.1.**) (Hofstede, 2001; Hofstede et al., 1990). Power distance refers to the power distribution (high score accepts hierarchy, low scores desires equality), individualism refers to the strength to which an individual has ties to a community (high score favours individualism, low score favours community), masculinity refers to how society views gender roles (high score favours traditional gender roles, low score favours equal gender roles), uncertainty avoidance refers to the degree to which ambiguity or change are tolerated (high score indicates rigid codes of belief and behaviour, a low score indicates a more relaxed attitude to deviations), long-term orientation refers to how society values long-term traditions and values (high score favours tradition, low scores favours innovation) and indulgence refers to the extent to which desires are controlled (high score indicates self-control and restraint, low score indicates gratification of desires) (Hofstede, 2001, 2011; Hofstede et al., 2010).

Similarly, the GLOBE studies (2004, 2007, 2014) quantitatively surveyed thousands of managers, at various levels and in a range of organisations, in more than sixty countries, on societal culture and leadership (Chokkar, Brodbeck, & House, 2007; House, Dorfman, Javidan, Hanges, & de Luque, 2014; House, Hanges, Javidan, Dorfman, & Gupta, 2004). Whereas in Hofstede's model

each country is given a profile of scores, in the GLOBE studies cultural groupings were created, representing common and similar profiles: Anglo, Eastern European, Germanic Europe, Nordic Europe, Latin Europe, Latin America, Sub-Saharan Africa, Middle East, Southern Asia and Confucian Asia (see **Table 4.1.**).

The GLOBE studies expanded on Hofstede's six dimensions of culture to performance orientation, assertiveness, future orientation, humane orientation, institutional collectivism, in-group collectivism, gender egalitarianism, power distance and uncertainty avoidance. Performance orientation refers to the degree to which a collective encourages and rewards group performance, assertiveness speaks to the level of assertion or aggression of an individual in their relationships, future orientation denotes the extent to which individuals plan for the future, human orientation indicates the degree to which a collective encourages fair, kind and altruistic behaviour, institutional collectivism refers to the level to which society rewards collective distribution of resources and collective action, in-group collectivism speaks to the amount individuals express pride and loyalty to their collectives (families, organisations), gender egalitarianism speaks to how much a collective minimises gender inequality, power distance indicates the extent to which a collective accepts authority and uncertainty avoidance refers to the degree to which society relies and holds on to norms and rules that alleviate unpredictable future events.

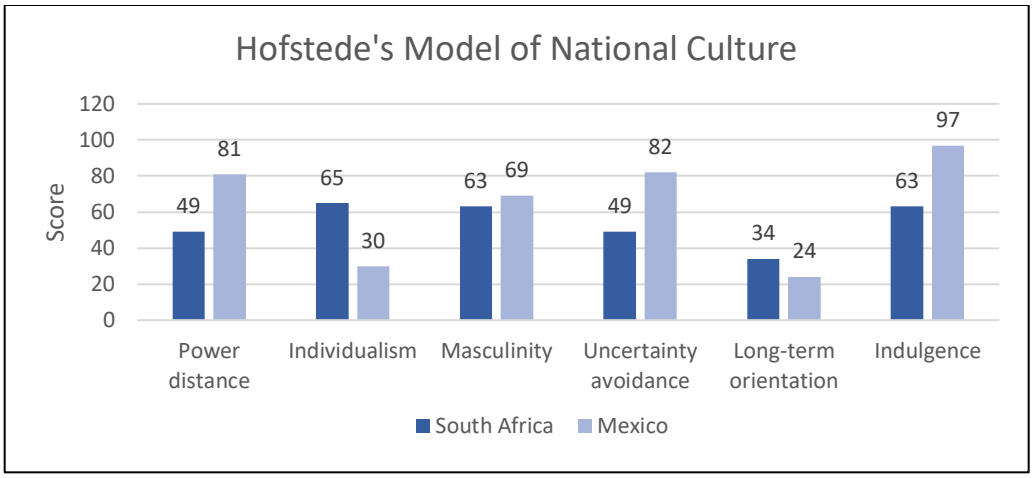


Figure 4.1.: According to Hofstede's Model of National Culture, South Africa and Mexico, both "Southern", "developing" and "resource-constrained" countries, are also culturally diverse (Hofstede, 2001).

Table 4.1.: An overview of the organisational and national cultural Hofstede and GLOBE country profiles for South Africa and Mexico and Indonesia (Hofstede, 2001, 2011).

Hofstede country profiles	
South Africa	**Mexico**
• Acceptance of hierarchical order. • Individualistic society (white population; caring for self and immediate family only). • Masculine society (assertive, decisive). • Tolerates change. • Short-term orientated (focused on traditions and not pragmatics or the future). • Culture of indulgence.	• Hierarchical society. • Collectivist society (strongly values community, committed and loyal to the in-groups to which they belong). • Masculine society (live to work, competitive). • Avoids uncertainty to a great degree (desires strict rules and legal systems to order or structure life, intolerant of unorthodox views). • Short-term orientated (heavy on traditions, but do not think of the future) • Highly indulgent (enjoy life).
Globe country profiles	
Sub-Saharan Africa **(contains South Africa)**	**Latin America** **(contains Mexico)**
• Relatively high scores for in-group collectivism. • Relatively low scores for power distance and gender egalitarianism. • Middle-range scores for the rest. • This means is that generally Sub-Saharan Africans are family or community orientated but that there is an unequal distribution of power and gender inequality. Interestingly, the value scores (the ideal or what society believes should be) are different to the practiced scores (actual or observed), for example, a higher gender egalitarianism score, indicating that they desire a more gender equality.	• Relatively high scores for in-group collectivism and power distance, suggesting a pride and loyalty to family and organisations and accepted authoritarian hierarchies. • Similarly, to Sub-Saharan Africa, there is a difference between the value and practice scores; such as a lower power distance score, thus they would prefer power to be more equally distributed.

There has been debate between the Hofstede and GLOBE studies on various elements (Hofstede, 2006, 2010; Javidan et al., 2006) yet both models can be critiqued for using individual-level data to make inferences at a collective, national-level, as well as a lack of clarity on the use and importance of national economic wealth in how it relates to various aspects of culture (Javidan et al., 2006; Smith, 2006). Critically, when using a model it must be stated that, "No single researcher or research team should own the cross-cultural research field," and be acknowledged that assumptions are made within all models, as none exhaustively identify all cultural constructs (Javidan et al., 2006). Indeed, oversimplifications are made and culture becomes rigidly categorised into a limited number of constructs, and the assumed stability in measurement scores over time, in contrast to reality, group spaces are complex, flexible, negotiated and continually developing – which stands in stark opposition to both the Hofstede and GLOBE models (Signorini et al., 2009; Tung & Verbeke, 2010). Yet the Hofstede and GLOBE models may still prove useful in providing commentary and possible explanations for lecturer assessment behaviour in different organisational and national cultural settings in this study and they do make the point that countries in a Southern context are quite diverse culturally among themselves; and these cultural differences have been shown to be important in terms of translating theory and practice from one culturally different context to another; it may or may not work, or require changes, due to cultural differences (Suhoyo et al., 2017; Suhoyo et al., 2014).

In this study, it is important to note that for the scores ascribed to South Africa, these measurements were taken during Apartheid and thus represent only the white minority and thus could differ significantly in reality, and so, the more contemporary GLOBE scores may be of greater accuracy and use. However, both models were consulted briefly during the analysis of the data collected.

Table 4.2.: Contextual diversity of sampling sites in the global South: Profiles of developing and resource-constrained countries in which data collection took place (The World Bank, The World Factbook – CIA).

Country	Population (Density)	Major Ethnicity	Colonial History	Major Religion/s	Language/s
South Africa	56.7 million (46.8 people per km^2)	Black African (Xhosa, Zulu, other) 80.2% European ancestry 8.4% Coloured (mixed ancestry) 8.8 % Indian/ Asian ancestry 2.5%	Netherlands/ Holland, formation of the union of South Africa in 1910 (Dutch & British rule), followed by Apartheid (racial segregation) & democracy in 1994	Christianity (≈ 85%) Islam (≈ 1.5%) Indigenous beliefs	12 official languages including English, Afrikaans, Sesotho, Setswana, Xhosa and Zulu
Mexico	129.2 million (66.4 people per km^2)	Ameridian-Spanish (60%), Ameridian (30%) and Other/ European (10%)	Spain, independence in 1810, "democracy" since 1917, "true" democracy since 2000	Catholic/ Christianity (≈ 90%)	Spanish (90+%)

Note: as these cultural profiles and contextual diversity descriptors have been represent "Northern theory" (performed or collected by researchers from the North, in the North and generally from a Northern perspective) and so a critical stance should be adopted when consulting these "Northern conceptualisations".

4.2. Interview Questions:

Prior to beginning interview:

- Introduce self (PhD student). Thank-you for your time.
- Go over consent form, any questions, must be signed and returned.
- Before we get to the interview, a little background (focus of interview): this forms part of my PhD research investigating how lecturers think about and practice assessment, and what factors influence their assessment practice, and hence I will be interviewing lecturers (course convenors of senior clinical rotations) in diverse, resource-constrained, "Southern" settings.
- Interview will be audio-recorded, transcribed verbatim, analysed, securely stored, then destroyed after a few years. Anonymity and confidentiality are ensured.
- Note taking during the interview for clarification and follow-up questions if necessary.
- Start recording.

A) Conceptions:

Purpose:

- What do you understand by the term "assessment"? How would you define "assessment"?
- What do you think is the purpose (reason for, goal/role of) assessment?
- Why do you practice assessment? What is it for? What does it achieve?
- Selection/judgement/decision-making? Summative?
- Certification/quality control/suitability? Fitness for practice/good doctor?
- Direction/diagnostic? Formative?
- Motivation/guidance? Feedback? Drive learning?
- Responsibility for society?
- Develop life skills?
- Punish students?
- Consequence if assessment done badly? Outcomes of assessment?

Learning effect:

- What role do you think assessment plays in driving student learning?
- Do you think that assessment has any impact on student learning?
- Positive/negative effect?
- What? How does it impact on student learning (study strategy/choices, quality)? Why?
- None/little? Forces learning? Short-term?
- Drives learning? Would they learn more/less if you took assessment away?
- Backwash effect of assessment? Do certain types of assessment affect student behaviour in different ways? How/why?
- Difference between junior/senior/UG/PG?
- Difference in how students learn from past papers? (Content vs comprehension/direct study behaviour)
- What else do you think drives student learning (besides/other than assessment)? Do lectures drive student learning more than assessment?

Criteria for soundness:

- How would you define a "good" or a "bad" assessment? What are characteristics of a good assessment? Why?
- What do you think makes up a good (high quality, technically sound) assessment?
- How do you determine the quality (technical and psychometric) of an assessment/question paper? (Statistical analysis)
- Are assessments checked/analysed before/after assessments?
- Have you ever been called in to give a reason for your marks/grades?
- What do you do when you act as an external examiner? Expectations?
- Validity? Sensitivity?
- Marking/grading – what do they represent/mean/tell us? Adjustment?
- Standard setting – relative/norm or absolute/criterion referencing? Pass/fail standard/50%?

- Rubric usage?
- Statistical analysis? Before/afterwards?
- Moderation?
- Curriculum coherence/blueprinting/alignment?
- Academic enabler? Use personal relationships/student exposure/experiences? (Subjective versus objective)
- Greatest shortcomings of your assessment system/practice?

Practice:

- How do you practice assessment? How do you go about setting/designing your assessments? Describe your assessment practice?
- Prompts/cues/clues/emphasis?
- Format/methods/tools? Question type? Coverage/exposure/sampling?
- Database/question bank?
- When? How often? (Frequency of testing, longitudinal/continuous/programmatic)
- Summative versus formative? Feedback?
- Relevance? Creativity? Innovation?
- Why? Reasons for practice?
- What has been your experience of assessment?

B) Influencing factors:

- What factors influence/direct/impact on your assessment practice?
- Personal? (agency, education, experience, motivation, reward, prestige)
- Interpersonal? (colleagues)
- Institutional/organisational? (reputation, culture, faculty guidelines/rules/policy/standards)
- Contextual? (time, resources/finances, workload/research/pressures, history, politics)
- Do you have to adapt your assessment practice? Why?

C) Demographics:

- What is your academic title/position? (Professor, senior lecturer)
- What is your discipline/course? (year/theory/clinical)
- What is your education/background/training (HPE experience?)
- How long have you been involved in assessment specifically? (years)
- What are your assessment responsibilities? (designing, marking, etc.)

End of interview:

- Stop recording.
- Thank-you for your time.
- Follow-up questions/interviews?

4.3. Reflexivity: Declaration of self

"How does who I am, who I have been, who I think I am, and how I feel affect data collection and analysis?" (Pillow, 2003).

I am a scientist. I have positivist tendencies. I once looked down upon qualitative research as "soft" and "easy". Much has changed! I made the jump from basic science, working as a medical cell biologist in a laboratory, to applied social science and education due to a deeply felt need and passion for education in South Africa. Whenever I feel tired or overwhelmed or disillusioned, I remind myself of my mission, my calling.

One of the other reasons I left laboratory-based research was a lack of translation of theory to practice. This was incredibly disheartening. I wanted to make a difference – or, at least, contribute in any, small and meaningful way. I have had to remind myself that the primary purpose of a PhD is for not for practical application, but the development or contribution to theory. I have had to temper my impatience and adjust my thinking. Rigour, evidence-based and theory-informed practice is now my motto when it comes to research.

I am a student, a researcher, someone who is constantly learning, asking questions and seeking answers. I do not yet see myself as an educationalist, because, the more that I study, the more I realise how little I know. It is humbling. It is scary.

Yet, I never realised how much I had grown over the last three years until I went to my first international conference in August 2019 (AMEE). I was surprised at my reaction and response to the conference, presentations and participants. I was offended, discouraged and saddened at the lack of real diversity and representation from the global South. I felt like a lone stranger in a sea of sameness, struggling against the tide. When I did speak up in sessions it felt like my "Southern" perspective was not understood or heard. It was through this that I realised that I have positioned myself in the field. I have placed my stake in the ground. While I still have much (!) to learn, I have

greater clarity on what I believe and what I stand for in Health Professionals Education and I am I feel driven to pursue it, believing that it is relevant, significant and needed.

I am a female. I am young. I am South African. I am a feminist. I am a Southern Theorist. I know what it feels like to be marginalised, to be silenced, to be spoken over. But I am also white, which gives me great privilege. All these experiences have made me hyper-aware of power and its imbalances. I need to be critically aware of my feelings (anger, hurt, frustration, irritation, hopelessness) towards the "West" and "North", "whiteness", males and "the patriarchy", that I do not jump to conclusions to judge literature or participants beforehand.

> *"I believe in Christianity as I believe that the sun has risen: not only because I see it, but because by it I see everything else" (C.S. Lewis).*

I am a Christian. This is the foundation for my worldview. It is why I care and am involved in education. It is why I desire to use my gifts and abilities to contribute to meaningful research. I is why I want to make an impact. It is why am a feminist and social justice advocate. It is why I am to lift the voices of the oppressed. It is not for me, it is not even for others (even though it hopefully will benefit them – as we are called as Christians to do), but for the glory of God.

4.4. Informed Consent Form

(HREC study approval: 689/2017)

The Process of Informed Consent

Academic staff who wish to participate in this study will be provided with information and consent forms (below) by email prior to their participation, allowing time for reading, discussion, consultation and decision-making. Prior to the interview the consent form will be explained to the participant in a private location, ensuring understanding before participating. At any point participants have the opportunity to query or withdraw from the process without fear of prejudice or penalty. Participants must have the capacity to consent, comprehend the information and agree with the process before being asked to sign the consent form in the presence of the researcher.

These consent forms are in English as all participating academic staff are expected to be able to communicate in English.

What is the title of this study?

Factors influencing lecturer assessment practice in diverse Southern contexts.

What is this study about?

This study will explore lecturers' conceptions (thoughts, views, beliefs) of assessment and what factors influence their assessment practices in final year/s medical programmes in a diverse range of Southern Health Professional Education contexts. The eventual goal of this study is to change lecturer assessment practice through training and professional development interventions, but before this may take place, lecturer assessment behaviour first needs to be

understood. Therefore, interviews will be conducted with lecturers exploring what personal and contextual factors influence their assessment practice. The outcomes of this study will be used to develop a model that describes, explains and predicts lecturer assessment behaviour, that may then be used to inform the design of lecturer training and professional development programmes.

The reason for this study is that assessment has serious consequences in the Health Professional Education context; assessment in final year/s medical programmes determines whether or not students will graduate and eventually license as a medical practitioner, which in turn impacts on patient care and public safety. Therefore, assessment must be used to influence student learning in desirable ways; assessment opportunities must be learning opportunities. Thus, this study seeks to understand and improve lecturer assessment practices to ultimately better student learning and enhance patient care and public safety.

Who is eligible to take part in this study?

Participants must be final year/s medical programme course convenors, in any discipline (psychiatry, paediatrics, surgery, obstetrics and gynecology, anesthesiology, physiotherapy, audiology, etc.), at settings in the South.

What must you do to take part in this study?

One-on-one, face-to-face semi-structured interviews will be conducted with participants in a selected, safe and secure location arranged around the participants' time schedule. Interviews will be recorded, transcribed and securely stored.

A cyclic data collection approach will be used in which single interviews with individuals will be conducted in a number of cycles. In total 1-3 interviews may be conducted with participants over

a number of months (potential follow-ups dependent on what data emerges from other sampling sites), lasting 30-60 minutes per interview.

Are there any risks and/or benefits to taking part in this study?

There are potential benefits to taking part in this study; this study seeks to explore the factors influencing lecturer assessment practice, thus the findings of this study could potentially have a positive impact on lecturer training and professional development. In turn, this could beneficially impact student learning through improving assessment practices and ultimately lead to enhancing the quality of patient care, public safety and health care system efficiency.

The results from this study will be presented at national and international conferences, as well as be published in national and international peer-reviewed journals, and thus may have wider beneficial influences in other Higher Education and Health Professional Education contexts. Participants too will be informed as to the final results of this study in the form of presentations and publications; allowing the participants to benefit from their contribution to this study and making them aware of any possible interventions to employ in their own contexts.

The risks and discomfort associated with this study will be minimal, i.e. not greater than the harms and discomfort normally encountered in daily life. There may be some minor inconvenience associated with the contribution of participants' time and energy, but overall, the risk:benefit ratio of this study is favourable; low risk and high benefit.

What about confidentiality and wanting to stop participating?

Participation in this study is voluntary and participants may withdraw from the study at any time without any prejudice or penalty. All information gathered will be strictly confidential and anonymous.

All documentation surrounding this study will remain within the possession of the researcher alone. Electronic data that references participants' identity will only be accessible by the researcher and data will be stored on password protected computers. At no point will participant name or personal details be disclosed to anyone other than those directly involved in the research. All participants' identity will remain anonymous and be referenced by way of a pseudonym. Given the nature of the questions that participants will be asked, it is considered unlikely that responses given participants could lead to their identification. Should responses made by participants be quoted in any report-back or publication, the researchers will remove any content (noun, relationship, reference, etc.) that could lead to the identification of participants by readers. Even were any participant to be identifiable, the content of the interviews is such that its revelation would not place participants at any risk.

Once the required reports have been generated the personal information of participants will be destroyed and the content of the information stored by means of pseudonyms. This data will be stored for approximately five years, in case there should be any queries on the research.

Will there be reimbursement for participation?

No. The interviews will take place at a location of the participant's choosing and arranged around the participant's time schedule; therefore, the participant should not incur any travel or other costs. Therefore, contribution of participant's time will not be financially compensated.

Any further questions?

Should participants have any questions with regards to the study, please email the researcher on smtdan013@myuct.ac.za.

Consent Form

I, _____, fully understand the study and consent to taking part in it.

(Print name)

Signature:_____

Date:_____

Table 5.1.: Pilot study: Lecturers' Conceptions of Assessment from South Africa, with illustrative quotes.

Outcome Space	Undirected	Content-focused/ Reproduction-directed	Competency and Conceptually-focused/ Application-directed
Purpose	• None: serves no meaningful purpose • Negative: an irritation or necessary evil *"A necessary … irritation" (SA3)* *"A necessary evil" (SA6)* *"[Assessment] has no real meaning for the student in terms of the following year" (SA6)*	• Summative: measures factual recall/reproduction of knowledge • Formative: forces/motivates learning • Academic focus & teacher-centred: content-expert > student conceptual development • Gatekeeper: guard profession and protect reputation > student *"[That they have] grasped the basic ideas" (SA12)* *"Motivator to get people studying" (SA7)* *"What the lecturer thinks is important" (SA12)* *"I am a representative of the profession" (SA7)*	• Formative focus: learning and student development/ conceptual change = student-centred • Social and moral focus: patient and society-centred *"Be a safe … doctor" (SA1)* *"Responsibility towards society" (SA12)* *"Responsible to civil society … good practitioners" (SA8)* *"Assessment, in the end, has to be a reflection of … how the student has grown in terms of that subject field; how he has developed" (SA4)*

Learning effect	• None: has little meaningful effect on learning • Negative driver of learning *"A bit of fright … fear is better … a bit of a scare … get their stress levels up … anxious" (SA1)* *"The threat of a test" (SA6)*	• Motivates learning: but directs undesirable learning behaviours (superficial learning & a short-term perspective) • Students focus on marks/passing *"[Testing the] minimum level of factual knowledge" (SU3)* *"Short-term knowledge and short-term recall" (SA11)* *"Very little long-term retention" (SA1)*	• Positive driver of learning: directs desirable learning behaviours (application of knowledge, deeper understanding & a long-term focus) • Lecturers focus on student development & clinical competency *"Formative thing" (SA7)* *"Insight … deeper level" (SA4)* *"Retain for as long as possible" (SA9)* *"Feedback" (SA2 & SA4)*
Technical knowledge	• None: little critical thought given to their assessment practice • Poor or misconceptions: limited or incorrect understanding of technical aspects of sound or high-quality assessment practice *"[Standard setting is a] thumb suck" (SA6 & SA12)* *"I just compensated by arbitrarily … by hiking up their marks … The*	• Simplistic: still developing • Expressed through: aware of limitations and shortcomings (subjective practice), acknowledge uncertainties & concerns (desire for sound and high-quality assessment practice) *"It's a very subjective thing … I have to say it's a gut feeling … There is no objective or*	• Sophisticated understanding: expresses concepts of validity, reliability, educational impact, practicality and acceptability • Demonstrated through: critique of unsound assessment practice and use of standardised assessment sheets/ rubrics, blueprinting/ alignment and continuous assessment *"Standardised assessment sheet" (SA2)*

	assumption that the average mark would be somewhere around about sixty-five, and of course that's a complete assumption, [but] if you see marks that are substantially less than that then you can adjust those or not as you see fit" (SA6)	*quantitative measuring instrument I use" (SA3)* *"I hope we don't make mistakes that often, but I have a sinking feeling that we do make them quite often … We make mistakes" (SA2)*	*"Rubric" (SA9)* *"[Assessment] content has to be relevant" (SA4)* *"Patient population" (SA8)* *"Continuous assessment" (SA2 and SA4)*

Table 6.1.: Second phase: Lecturers' Conceptions of Assessment from a second South African context, with illustrative quotes.

	Detached practitioner	Emerging equilibrium	Engaged educator
Purpose	• Assessment drives learning (summative – administrative) • Clinical competency goal (more of a content-expert focus) *"It is just because we are required to assess them as part of their mark" (SA13)* *"[Assessment must] motivate them [to study], because if they don't get a pass then they have to repeat the block" (SA14)* *"If the student fails the exam, it is quite an administrative nightmare … So, it is an incentive to not let them fail and if they are really bad you'll perhaps try and find a way to give them at least 50%" (SA13)* *"The goal of assessment is to determine who actually has [disciplinary] knowledge, who has good clinical skills … make sure that there is a certain*	• Assessment drives learning (both: formative and summative) *"If you didn't have an exam, then people might not work as hard, so that's one concept of what assessment driving learning … There is certainly a group of people who are probably going to spend more time with their books and learning because there is an exam … The other thing is, when one understands what is coming in an assessment, then that also drives how you learn … So, because students know there is a clinical exam with case vignettes and are expected to make a diagnosis and an investigation plan and a treatment plan, then they start to think like that … and they will apply*	• Assessment drives learning (formative) • Student learning and development goal • Clinical competency goal (more of a moral/social focus) *"Formative case presentations … Feedback … how they can improve and it doesn't count for a mark … It is purely for the student's benefit, there is no mark … the students are encouraged to use that as an opportunity to learn" (SA14)* *"A clinical scenario … a real-life patient … how to manage a patient, treat them properly … the scope of practice of an intern" (SA14)* *"Because students know there is a clinical exam with case vignettes and are expected to make a diagnosis and an investigation plan and a treatment plan, then they start*

	standard of knowledge or certain level of knowledge and clinical abilities. So the assessment of that knowledge and their clinical skills is what we are trying to achieve with the end of block examination" (SA16) Concrete: practical task (administrative – performed in order to get a mark for each student during the block before they can be passed off to the next block) *"Our main job is not doing assessment, it is running the ward and doing other things, and it [assessment] is sometimes seen as a chore we have to do" (SA13)* *"It is just because we are required to assess them as part of their mark" (SA13)* *"They [students] need to pass that exam and then their overall mark must also be a pass" (SA14)*	*themselves in that way when they are exposed to that in their clinical environment. And I think that is probably more important when you talk about assessment driving learning" (SA18)* Both: concrete (administrative) and abstract (learning) *"It's a high stakes event … failing an examination has massive implications. So, people fear that and so they learn, they study, so that they pass their exam. If you didn't have an exam, then people might not work as hard, so that's one concept of what assessment driving learning … When one understands what is coming in an assessment, then that also drives how you learn … and they will apply themselves in that way when they are exposed to that in their clinical environment" (SA18)*	*to think like that … and they will apply themselves in that way when they are exposed to that in their clinical environment" (SA18)* Abstract: learning and developmental focus *"Reflection on their professional and personal development … promote agency … as future interns … do they feel prepared, energized, inspired, an agent of change … They [are] welcomed as part of the health team … feeling like they are functioning as a doctor than just a student in this big hierarchy … That's as valuable for their formation as … professionals to practice as the formal feedback on the report" (SA17)*

Temporal perspective	Short-term (administrative task – block/ course focus) *"It is just because we are required to assess them as part of their mark" (SA13)* *"[Assessment must] motivate them [to study], because if they don't get a pass then they have to repeat the block" (SA14)*	Both	Long-term (formative development and clinical competency)
Assessment literacy	• Poor or incorrect • Tacit/implicit craft knowledge: content/ disciplinary experience & expertise *"It's mainly just from our own experiences [as clinicians] … They have examined before … [Some] are quite good at actually doing this well … Every examiner may do it differently" (SA13)* *"Most of the examiners … are also just clinicians … quite seasoned examiners … so experienced" (SA14)* *"The examination is often not standardised, because you have different clinicians, different individuals who run the portfolio exam … The examiners*	Both: tacit knowledge (as a clinical-expert) and HPE evidence-based practice are needed/important *"I can never remember the terms, formative and…[summative]? I can never remember and I always get confused about which is which" (SA18)* *"So that kind of training and ongoing capacity development around assessment … I think more of that could be happening. You know, you think of new staff coming in and need to design an assessment … it's quite challenging setting exam papers or … designing assessment rubrics … When I was designing exam*	Advanced/explicit assessment knowledge and HPE evidence-based practice *"Each of the [assessment] formats does sort of assess different things and different aspects … There is no point having multiple assessment points but they are all the same; you need to decide what it is you are testing and focus your assessment on that … So, in other words, are you testing what you want to test?" (SA18)* *"I blueprint everything" (SA15)*

	are all clinicians … experienced clinicians" (SA16) *"[Our assessing clinicians] are experienced at examining at both the undergraduate and postgraduate levels … All of our examiners are generally experienced" (SA18)*	*papers … it was mostly from my own prior reading and learning; and then one just develops skills over time" (SA17)*	
Identity	Clinician *"I'm a clinician, I am not [a teacher]" (SA14)* *"I am a clinician first of all" (SA16)*	Both: clinician and educator	Educator *"Formative case presentations … Feedback … how they can improve and it doesn't count for a mark … It is purely for the student's benefit, there is no mark … the students are encouraged to use that as an opportunity to learn" (SA14)*
Role	• Role: administrative, manager, operator (pragmatic or mechanistic view: simply implements, no ownership) • Clinical content/disciplinary expert *"I just got involved [in assessment] because it fell [into my lap], I was delegated to be the course convenor … I inherited it … I am not sure exactly where he got it from … I don't really know where it comes from" (SA13)*	• Role: developing (negotiation/ tension/ compromise) • Balance content/ disciplinary expertise with educational training *"The vast majority of convenors and examiners in the clinical years are clinicians … The vast majority of the teachers are first and foremost clinicians and their teaching*	• Role: educator and assessor (initiates and in control/ownership of assessment practice) • HPE evidence-based practice *"I … do most of it myself" (SA17)* *"When I inherited this programme … they only had one formal assessment, which was the observed consultation … And then when I started in*

	"We just rotate through the department whose turn it is to examine" (SA13) *"As the organiser for the students … I would assume [it] would be someone else's responsibility" (SA13)* *"I tend do to the sort of managing the students on the ground, a lot of these nuts and bolts of the programme" (SA16)* "*I catered it to what I felt was appropriate for surgery [and] I based it on other disciplines … A lot of the exam is structured on … the surgical college exam" (SA18)*	*commitments are added on top" (SA18)*	*2011 we evolved to the OSCE" (SA15)* *"Yes, so I designed it [a rubric] … I do most of it myself … here are quite a lot of changes I made initially" (SA18)* *"I blueprint everything" (SA15)*

Table 6.2.: Final Outcome Space describing Lecturers' Conceptions of Assessment from diverse Southern contexts, with illustrative quotes.

	Passive operator	**Awakening inquirer**	**Active owner**	**Scholar**
Purpose of assessment	Administrative (a task, to get marks) *"Evaluation still has to count towards a goal, it must still provide a mark" (SA1)*	Psychometric (to measure student learning) *"Evaluation is a way of proving that the students have gone through the work" (SA1)* *"[Testing the] minimum level of factual knowledge" (SA3)*	Moral & Social (student learning, public service & patient care) *"Whether or not he is informed enough fo r… practice one day" (SA5)* *"We have to take into account every single patient, and it's a patient of the students'" (MX4)* *"Are they able to connect the classroom knowledge to the real-life setting?" (MX13)*	
	Summative (marks for decision making)			
		Formative (marks as feedback) *"We do feedback after the exam – what marks they actually got … so, that they can address any issues right there than come back months later when no one can actually remember r... Because I am the convenor, I do get the occasional student sending me an email or coming to me, that the student is upset, that they should have gotten a higher mark, and then*	Formative (deliberate feedback for student development) *"Formative" (SA7, SA14)* *"Feedback … how they can improve … It is purely for the student's benefit … an opportunity to learn" (SA14)* *"That constant dialogue with a student is a feedback, it is not a grade, it is a feedback" (MX4)*	Formative (student learning, teacher performance, curriculum appropriateness, practice preparedness) *"We do review our programme every year … we look at student feedback, which I think it important, and we try to improve" (SA15)*

	you are trying to sort that out months later is a bit of a mission because nobody can remember much about what happened. So, it is only this year that we decided to give students feedback immediately and then deal with any issues straight away" (SA13)	*"I need to know if they are doing correct things or not … it's essential that we give them feedback … In every moment that we can" (MX8)* *"You try to identify in yourself which areas are you weak, that you are going to learn about … that you are going to take action in" (MX11)* *"They need to do a reflection … at the end of their rotation, talking about how they think their development was, related with the whole bunch of outcomes … [The] reflection is the one of the most important ones … the most valuable thing is that they need to make these reflections and … to say, "What did you do? What did you get? What is needed? And what are you going to do for that?"" (MX12)*	
Concrete and Practical (a task) *"A chore we have to do" (SA13)*			

"So, students come through, we just assess them and throw them out" (SA8) *"Educate a person to be a doctor, to think as a doctor and act like a doctor, but also to take an exam" (MX4)* *"I would be very, very focused on trying to get my students pass any test" (MX9)* *"They come with me, and then you do like as a subjective evaluation, at the end you just check the list … Lots of the time it's very numerical … In the end, we have to assess students" (MX13)*			
	Abstract (a tool for conceptual development) *"I believe it [assessment] is a tool to see if the student gets prepared or gets ready to do the objectives that we want him to have when he becomes a doctor" (MX2)* *"We're trying to make that the students think and not just memorise the correct answers" (MX8)* *"I'm more comfortable with saying that he's "doing okay", he's "proficient" or he's "lacking" – I like that way better" (MX13)*		
Administration (capture marks) *"We are required to assess them as part of their mark" (SA13)* *"I would be very, very focused on trying to*	Content reproduction focused *"Minimum level of factual knowledge" (SA3)* *"Make sure that there is a certain standard of knowledge or certain*	Competency development focused *"The certifying someone as competent to a particular task" (SA12)* *"Insight… deeper level… [Assessment] has to actually be in line to help*	All stages of learning and development (Miller's pyramid) *"Learning opportunity" (MX3)* *"Professional and personal development" (SA17)*

	get my students pass any test" (MX9) *"At the end of surgery he came "Evaluate me"; so they try to be and behave as well as they can, because they know they get evaluated every time … most of them are just worried because they want (good) grades" (MX10)* *"They come with me, and then you do like as a subjective evaluation, at the end you just check the list … Lots of the time it's very numerical … In the end, we have to assess students" (MX13)*	*level of knowledge" (SA16)*	*students improve their skills or their abilities to become better doctors" (SA4)* *"A tool to see if the student gets prepared … when he becomes a doctor" (MX2)* *"Competence, competence, competence" (MX6)*	*"[This HEI] does not prepare students to pass an exam; it's for them to be good practitioners and to be competent" (MX6)* *"A global vision of his hospital work" (MX10)* *"There are scales on "great performance", "he could do that better", "he needs to improve a lot", "he failed"" (MX12)* *"They try to find points for the knowledge, social skills and personal development" (MX13)*
	Local (individual student and course) *"So, students come through, we just assess them and throw them out" (SA8)* *"If they don't get a pass then they have to repeat the block" (SA14)*		Global (student learning over a programme) *"The development of the student as they grow … progress … improve" (MX1)* *"Assessment, in the end, has to be a reflection of …*	Global (student learning over a lifetime and HPE more broadly) *"Measuring student performance, but also the formative promotion of student learning, where the feedback helps to*

	"Educate a person to be a doctor, to think as a doctor and act like a doctor, but also to take an exam" (MX4) *"If they failed my course they have to do another year, an entire year of waiting; having a 12-week course but having to wait one year so you can rewrite! So, I think we can change the grades in defence of the students not achieving" (MX8)*	*how the student has grown in terms of that subject field; how he has developed" (SA4)* *"The first thing is that you'll have to define what the characteristics of ... what the level of knowledge is that one would expect of a first-year student, on that level, what is the level of knowledge of a second-year student, and what level of knowledge does a third-year student have, so one can determine a profile for each and then… you can set up your questions around that profile" (SA4)*	*improve … And then the function of evaluating, you know, one's course or the programme as a whole'' (SA17)* *"We do review our programme every year … we look at student feedback, which I think it important, and we try to improve" (SA15)* *"Programmatic assessment is our goal … Programmatic assessment as a guarantee, because of cycles of measuring, reflection, giving thought to the results and then making a change, like strategies" (MX6)* *"Assessment is going to try drive learning, but it also drives teaching … The assessments are the most important thing in medical education" (MX6)* *"I think assessment is quality control of what*

			we're doing; so it goes both ways it goes, "What is the students doing wrong?" and "What we are doing wrong?" … So I think assessment goes both ways – with the professors and students" (MX7) *"I do give them a lot of feedback, and I expect feedback from them" (MX13)* *"I'm going to give you the tools that you are able to evolve in your learning of medical practice. And that's done through a lifetime. If you stop learning you better quit" (MX13)* *"I think students should be responsible for their learning, and I'll say this, I'm thinking of me right now, as I'm 50 years old, and I'm still learning" (MX9)*

Temporal Perspective	Short-term (task, to get marks) *"[Assessment] is purely for a short-term goal and has no real meaning for the student in terms of the following year" (SA6)* *"They need to pass that exam" (SA14)* *"Educate a person to be a doctor, to think as a doctor and act like a doctor, but also to take an exam" (MX4)* *"They have to know their stuff by the following Friday … We are required to assess them as part of their mark" (SA13)*	Long-term (to develop competency for future clinical practice) *"To become a better medical student and a better physician" (MX3)* *"Be a safe … doctor" (SA1)* *"Long term it has to do with the certifying someone as competent to a particular task – the responsibility towards society" (SA12)* *"Seven competencies … So there are things that the student needs to learn how to do, professional activities, or to have these experiences" (MX6)* *"I do not give them a final exam; I don't believe in final exams; I think they not good… I think final exams are just a test of how much you can cram in your head… It is not even real life for a physician – every day is a test … What we should be trying to evaluate is how*	Long-term (to develop competency for future clinical practice and improve HPE practice) *"We do review our programme every year… we look at student feedback, which I think it important, and we try to improve" (SA15)* *"The challenge is how to help the students [learn] and how to help the professors to teach [and assess]" (MX6)* *"I'm going to give you the tools that you are able to evolve in your learning of medical practice. And that's done through a lifetime. If you stop learning you better quit" (MX13)*

		I'm able to solve that problem … Because it prepares you for real life. I believe in multiple assessment every day" (MX9)	
Assessment Literacy	Simplistic (unaware or a novice; implicit clinical craft knowledge and expertise) *"Our own [clinical] experiences" (SA13)* *"I like to think that by asking different students different things, I can gauge an overall impression of their place on the marks scale. However, I know it's certainly not something one can prove scientifically… It's a very subjective thing … I have to say it's a gut feeling … There is no objective or quantitative measuring instrument I use" (SA3)* *"I think everyone already knows this intuitively because most of the lecturers there [are] … long standing members [and] have a good understanding of it" (SA5)* *"I don't know if there is some objective way to assess the answers" (MX5)* *"It's a gut feeling one just has, one that I've picked up along the line … It's something that comes with experience and time, you get to know where the standard lies, but there isn't an instrument I can measure it with" (SA2)*	Developing (implicit clinical craft knowledge and basic HPE awareness – still needs to advance) *"I also realized is that I actually had very little idea of what to do … We actually had lectures about lecturers … It had value to me, but also frightened me a bit, in the sense that I knew I was lacking" (SA10)* *"We often assess incorrectly and because we don't ask higher order questions, we don't test deep knowledge, but rather we test surface knowledge which they can memorise very easily and can regurgitate" (SA4)* *"They know 70% of the knowledge, they are quite good; and if they know*	Emerging (more advanced HPE knowledge) *"We have a rubric we use … in the study guide there is a rubric that they see, which we use to evaluate them on" (SA2)* *"That mapping part I was talking earlier" (MX3)* *"You should be assessed by multiple people … You evaluate every single thing, and I think it's actually more … fair for the students … I believe in in multiple assessment and doing it a lot of times … I use a rubric" (MX9)* *"They all measure different things; like the*

			69.9 they are bad - so it's so arbitrary. That's why I'm saying that the final grade they have 69.9 - should I put 70? It's a very small part of their qualification. I don't know what it means. But in Mexico that's the system that we have … [Rather] just pass or fail? I don't know … You know or you don't know; you deserve it or you don't deserve it … That number doesn't makes you a better doctor … That number doesn't make you better or worse I think" (MX8)	*theory and then the app measures like the global behaviour of the students in different scenarios, and the portfolio is just like a record of activities" (MX10)*
Identity	Clinician *"I am a clinician first of all" (SA16)* *"I am very clinical … We're not teachers; we're clinical doctors that give classes" (MX13)*	Clinician-Educator *"I'm a professor, but I am a clinician" (MX4)* *"The role is mixed; balanced clinician and teacher; equal" (MX6)* *"A clinician trying to be a teacher" (MX8)* *"I find it really hard to say that I'm a teacher or a professor because my wife's grandmother, she's*	Educator *"Your role will be to advise students and support them" (SA1)*	Educationalist and researcher

		a teacher, but a teacher teacher, and she was one day like, "Ah, you should not call yourself a teacher because you did not study education and everything" and that really got into me and I'm like "I'm a doctor, not really a teacher, but if I can teach the things that I have because of my experience or age or whatever, I'm a teacher, but not a teacher teacher" (MX5) *"I am a clinician convinced that the way I practice needs to have education in sight… It's a two-way street; it works both ways… Finding that halfway thing is something I believe" (MX3)*		
Role and Responsibility	Administrator (manager, operator) and content-expert (clinical knowledge) *"Our main job is not doing assessment, it is running the ward" (SA13)*	Teacher as a content-expert and educator *"I'm going to ask you this therefore you must learn it" (SA6)* *"What the lecturer thinks is important" (SA12)*	Teacher as an educator (guide and mentor) *"Your role will be to advise students and support them" (SA1)* *"I think students should be responsible for their*	Scholar (performs educational research) *"There's like this research team that looks for the new tendencies of other schools who try to imitate or listen to*

	"I tend do to the sort of managing the students on the ground" (SA16) *"You have to remind me, "Doctor, can you evaluate my surgery?" otherwise your course coordinator can come and say, "You didn't do anything these 15 days, we have no record of activity of you"" (MX10)*	*"The professor … he's the expert, the teacher has to give the information" (MX4)* *"Experience by the doctor is still very important to transmit to the students; so, I think it's very important that they attend classes, and they hear the class from the from the professor … I think the interaction with the professor is very important" (MX7)*	*learning, and I'll say this, I'm thinking of me right now, as I'm 50 years old, and I'm still learning" (MX9)*	*them … schools have their own like self-assessment to see what works what was good or what was wrong" (MX6)*
Reflexivity	Poor and passive (unquestioning or noncritical acceptance) *"It is what has always been done, which doesn't necessarily have to be the best way to do it, but, it is sort of the way we do it now" (SA13)* *"I haven't really thought about it [about standard setting] long and hard" (SA2)*	Active and questioning (becoming critical) *"I hope we don't make mistakes that often, but I have a sinking feeling that we do make them quite often … We make mistakes" (SA2)* *"We actually had lectures about lecturers … that's how I got involved, and I think because one starts to ask questions … The first time we started thinking about assessment … we learned*	Critical and evidence-based practice (consumer of HPE, implement best practice principles) *"I have to say honestly that until a year ago I wasn't a big advocate of the whole story of continuous evaluation … I believed in giving a man a proper exam, forcing him to bite his nails and sit on his backside for a day or two to go through the work. Unfortunately, that is also how I studied … I have now developed*	Scholar (consumer and producer of HPE) *"There's like this research team that looks for the new tendencies of other schools who try to imitate or listen to them … schools have their own like self-assessment to see what works what was good or what was wrong" (MX6)* *"I can validate because there is evidence that*

		about different assessment methods and so there I started thinking" (SA10) *"Doctors, the other teachers, we don't talk about the test … We always send it to someone and they put it together, but we don't have a lot of introspective … we need that retrospection" (MX8)*	*other insights … I have come to be a big advocate of the whole concept of continuous evaluation" (SA1)* *"We have tried, we have jumped over models, different models, to see what it works … we will now have portfolios" (MX7)* *"I understand the portfolio may be something more robust … At the end of the clinical rotation, the students have about 30 evaluation forms of different aspects with different ratings … [The] portfolio gets everything; and I can have a better picture of my students" (MX12)*	*what we need to do something [new]. Sometimes we have been like the first pioneers in these kind of fields in Mexico; but there is evidence that we need to go over there" (MX12)*
Accountability	HEI (guided by rules and regulations, reputation/ standards) *"Well there's the tradition – we have a test and we have an exam" (SA9)*	Student (learning and development) *"[Our] formative assessments … We have to help them; we have to do something with them … And it usually results in a massive improvement" (SA7)* *"As a professor I can say, "Oh, it's your fault, you're not studying, it's your fault" but it's my responsibility we have to share. If a student, after he finishes in our school, doesn't get into residency, is a fail. It's a failure for his parents, and it's a failure for us, it's a failure for them. We have to see them*		

	"The University says it should be so" (SA6) *"The university's rules stipulate … We go about it the way we do purely because … the university recommends it" (SA1)* *"[We do] what we are expected to do in the faculty, they tell you, "This is what you are supposed to do" (SA15*) *"We have prescriptions … guidelines for that and you try to follow it … We are forced to take all of those elements into consideration … If you don't stick to the rules you'll always have a problem" (SA5)* *"Because they told us to, that we have to" (MX5)* *"The boss, the director; it was established from the*	*through. If you don't see it that way, we won't act, it's a shared responsibility between students and teachers… we are in the same boat" (MX4)* *"The emphasis shouldn't be on the promotion of the student, the emphasis should be on … has my student achieved what he had to achieve, and after that the promotion or lack thereof is an automatic choice" (SA2)* *"Say, "What is happening here…?" There is someone listening on a regular basis, someone taking an interest, right from the beginning … If students are regularly looked at … if you don't do well someone will chat to you and ask you, "What's going on here…?" Sort of examining the whole mentorship thing … We shouldn't wait until the end of the year when a student fails to actually know what's happening" (SA12)* *"I'm going to give you the tools that you are able to evolve in your learning of medical practice. And that's done through a lifetime. If you stop learning you better quit" (MX13)*

	beginning, they say, "This is this going to be like this" … because the boss says that" (MX8) *"It's given by the school" (MX13)*		
	Profession and Discipline (knowledge/content and reputation) *"I am a representative of the profession … I'm obviously protective over the company; I can't allow someone to go into the community, to sign off … and we let him go and then he kills people or he makes terrible errors in judgement … I mean that's accusation against us … We have a tremendous responsibility towards … the university and our profession" (SA7)*	Patient and Public (clinical competency; moral and social) *"There should be a minimum mark that is quite high on utterly basic concepts … important core knowledge … Because if they don't they will potentially do harm … to people" (SA6)* *"If [a student] doesn't succeed d… to be fair to the other students and to the community, we can't allow a guy or a lady like that to pass through the gate" (SA7)* *"Getting students to acquire the knowledge that they need to practice with … And I suppose in that sense it would safeguard society" (SA11)* *"Responsibility towards society" (SA12)* *"Responsible to civil society … good practitioners" (SA8)* *"We have [a] strong assessment, our quality control is high … [it's] a moral thing" (MX1)* *"I asked them to explain to me, "What if the patient is in the private hospital, in the public*

			hospital or in community services – like rural situations … [How] are [you] going to act at the three hospital levels hospital; in the private, public one or rural?" It's not the same answer. If they can answer questions like that, I can know, I can see if they are with the knowledge, with the preparation, with the skills to treat the patient" (MX11)
Emotional valence	Negative to neutral *"A hurdle they have to cross" (SA4)* *"A necessary evil" (SA6)* *"I also believe that without that bit of adrenaline pumping you don't really learn … Just a bit of a fright … A bit of fright … fear is better … a bit of a scare … get their stress levels up … anxious … You're not going to study if you're not the tiniest bit anxious" (SA1)*	Neutral	Positive *"Inspiring people to learn" (SA1)* *"What I find wonderful about [our assessment], that's really positive for me, is that they look at how we can help people" (SA7)* *"Change the assessment as it currently stands so that students can also see… the value… rather than have it be a process of frightening them into realising that… you will fail and you won't pass" (SA1)* *"I enjoy it" (MX9)* *"It's my own interest … I was really involved in medical education … I was fascinated by the outcomes, and curriculum design is one of my favourite parts … I was really fascinated with the idea about this outcome evaluation process, how we could really do it … to see really what is going on with this student is fascinating; how the portfolio could help us to do that better. So I like evaluation; that's why I was involved in it and I try to keep involved" (MX12)*

			"I do think that I enjoy [assessment]; so no, I don't take it like as a negative" (MX13)

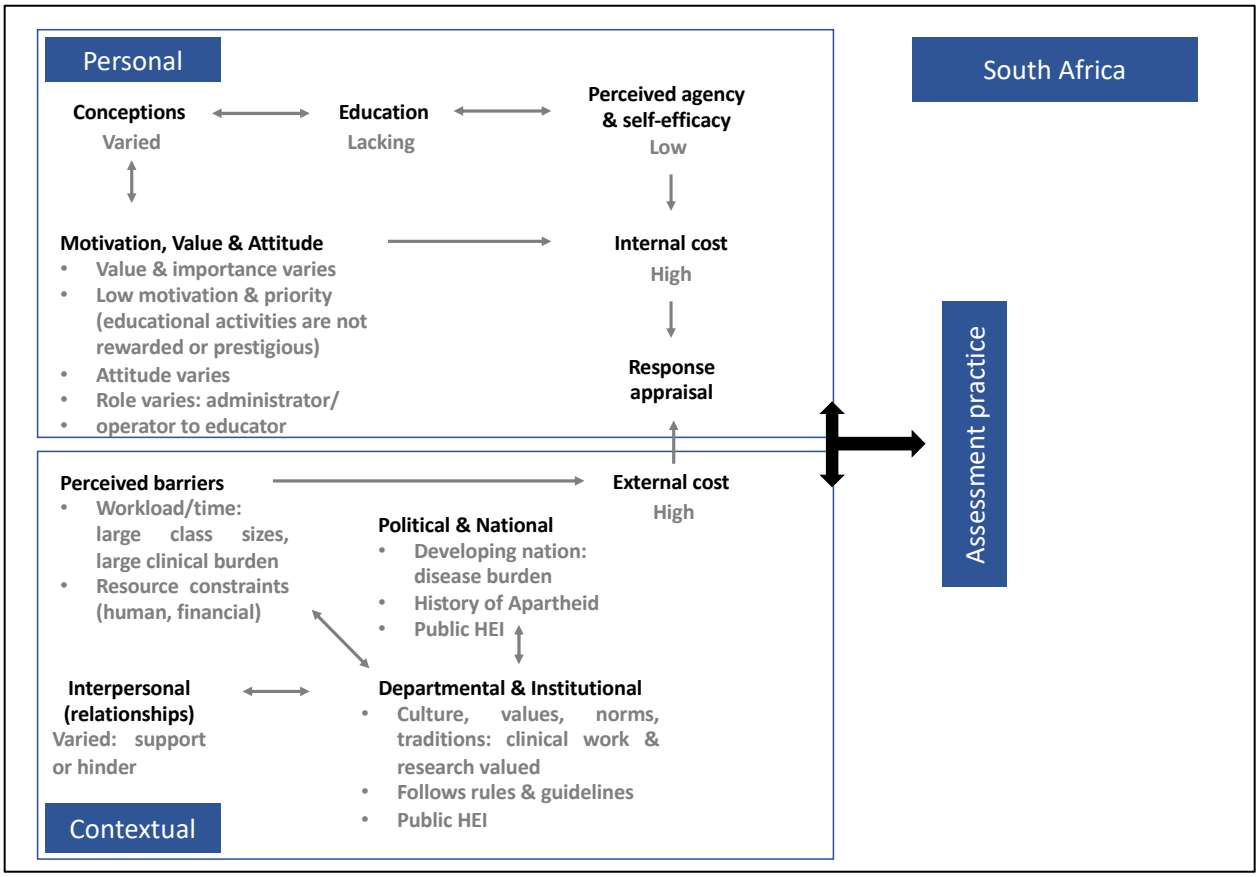


Figure 7.1.: A model of lecturer assessment practice in South Africa. A number of personal and contextual factors have been shown to influence lecturer assessment behaviour in South Africa. Of particular interest, is the national contextual factor of burden of disease (immense need for health care), which may account for the strong clinician identity and prioritising of clinical work, and the fact that all medical HEIs in South Africa are public (these lecturers experience time and resource-constraints in terms of their educational activities due to their clinical workload).

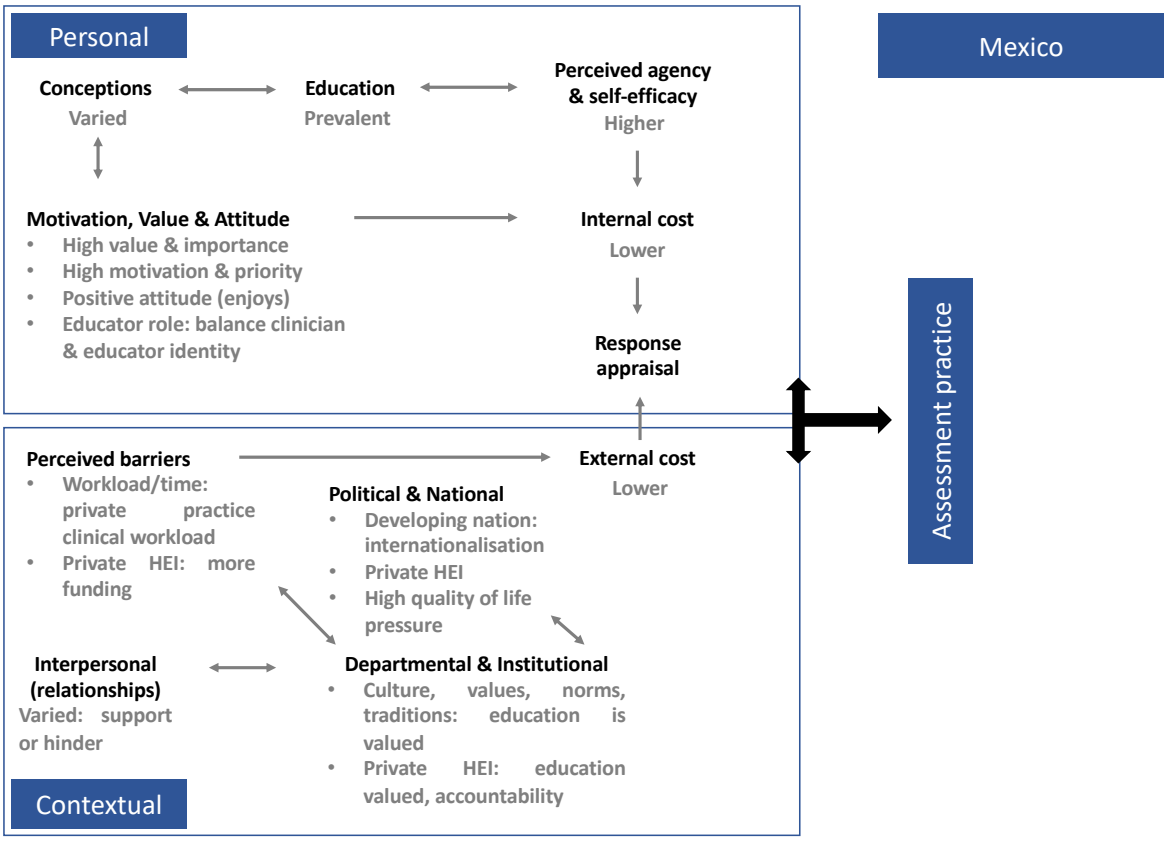
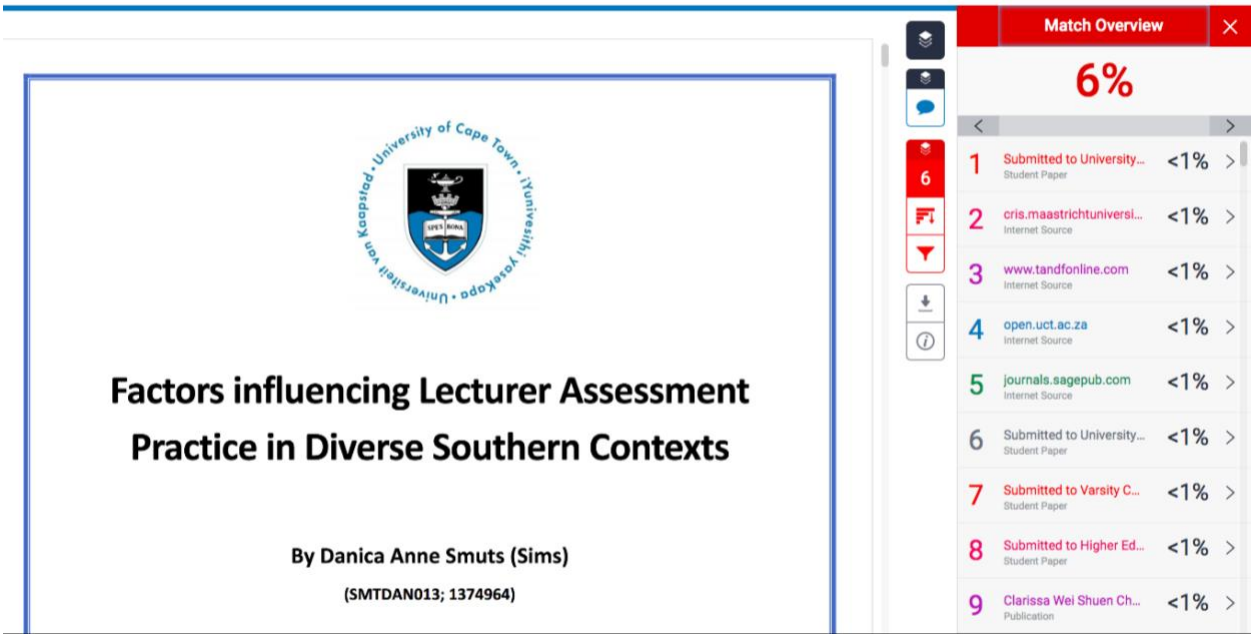


Figure 7.2.: A model of lecturer assessment practice in a Mexican context. A number of personal and contextual factors have been shown to influence lecturer assessment behaviour in a Mexican context. Of particular interest is the national and institutional factors of culture, privatisation of HEI, as well as the personal motivation for a high standard or quality of living. In Mexico, lecturers at a private medical school were sampled, where they undertook their private clinical practices at the private HEI hospital for money (their salary) but also performed educational activities there (teaching and assessment) due to a personal desire or motivation or enjoyment of education and students.

Turnitin Report



Factors influencing Lecturer Assessment Practice in Diverse Southern Contexts

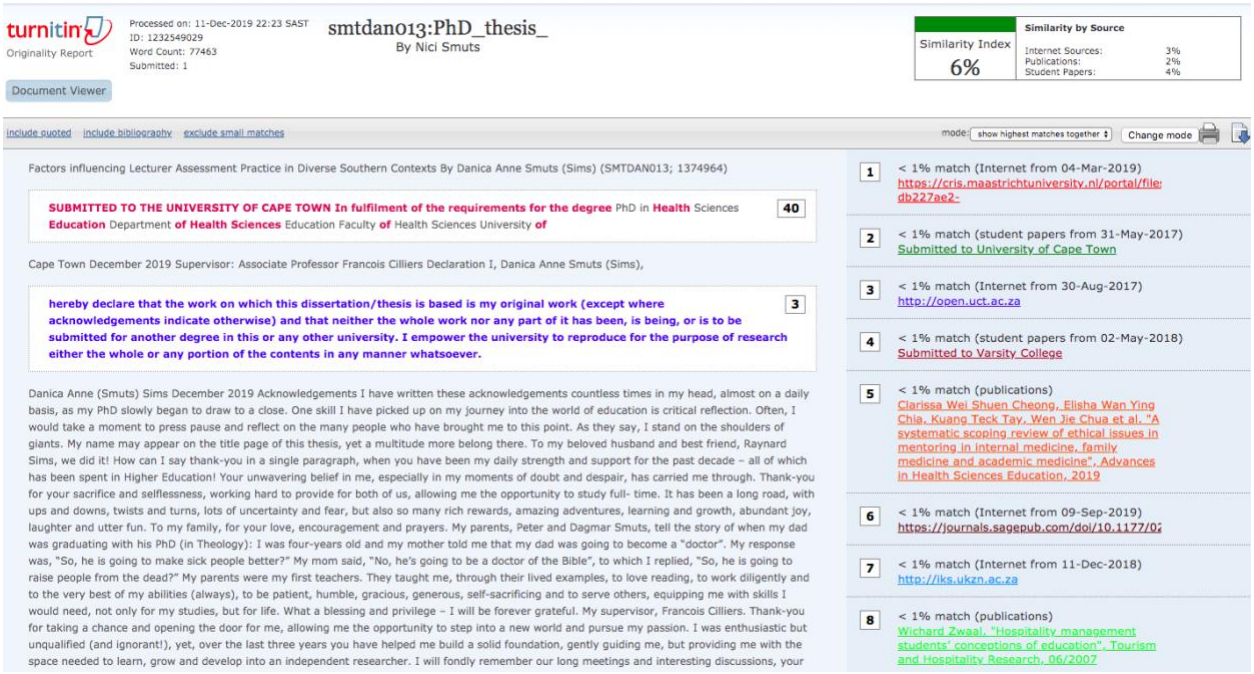
By Danica Anne Smuts (Sims)

(SMTDAN013; 1374964)

Match Overview

6%

1	Submitted to University... Student Paper	<1%
2	cris.maastrichtuniversi... Internet Source	<1%
3	www.tandfonline.com Internet Source	<1%
4	open.uct.ac.za Internet Source	<1%
5	journals.sagepub.com Internet Source	<1%
6	Submitted to University... Student Paper	<1%
7	Submitted to Varsity C... Student Paper	<1%
8	Submitted to Higher Ed... Student Paper	<1%
9	Clarissa Wei Shuen Ch... Publication	<1%



turnitin Originality Report

Processed on: 11-Dec-2019 22:23 SAST
ID: 1232549029
Word Count: 77463
Submitted: 1

smtdan013:PhD_thesis_
By Nici Smuts

Similarity Index 6%

Similarity by Source
Internet Sources: 3%
Publications: 2%
Student Papers: 4%

Document Viewer

include quoted include bibliography exclude small matches

mode: show highest matches together Change mode

Factors influencing Lecturer Assessment Practice in Diverse Southern Contexts By Danica Anne Smuts (Sims) (SMTDAN013; 1374964)

SUBMITTED TO THE UNIVERSITY OF CAPE TOWN In fulfilment of the requirements for the degree PhD in **Health** Sciences **Education** Department **of Health Sciences** Education Faculty **of** Health Sciences University **of** 40

Cape Town December 2019 Supervisor: Associate Professor Francois Cilliers Declaration I, Danica Anne Smuts (Sims),

hereby declare that the work on which this dissertation/thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university. I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever. 3

Danica Anne (Smuts) Sims December 2019 Acknowledgements I have written these acknowledgements countless times in my head, almost on a daily basis, as my PhD slowly began to draw to a close. One skill I have picked up on my journey into the world of education is critical reflection. Often, I would take a moment to press pause and reflect on the many people who have brought me to this point. As they say, I stand on the shoulders of giants. My name may appear on the title page of this thesis, yet a multitude more belong there. To my beloved husband and best friend, Raynard Sims, we did it! How can I say thank-you in a single paragraph, when you have been my daily strength and support for the past decade – all of which has been spent in Higher Education! Your unwavering belief in me, especially in my moments of doubt and despair, has carried me through. Thank-you for your sacrifice and selflessness, working hard to provide for both of us, allowing me the opportunity to study full- time. It has been a long road, with ups and downs, twists and turns, lots of uncertainty and fear, but also so many rich rewards, amazing adventures, learning and growth, abundant joy, laughter and utter fun. To my family, for your love, encouragement and prayers. My parents, Peter and Dagmar Smuts, tell the story of when my dad was graduating with his PhD (in Theology): I was four-years old and my mother told me that my dad was going to become a "doctor". My response was, "So, he is going to make sick people better?" My mom said, "No, he's going to be a doctor of the Bible", to which I replied, "So, he is going to raise people from the dead?" My parents were my first teachers. They taught me, through their lived examples, to love reading, to work diligently and to the very best of my abilities (always), to be patient, humble, gracious, generous, self-sacrificing and to serve others, equipping me with skills I would need, not only for my studies, but for life. What a blessing and privilege – I will be forever grateful. My supervisor, Francois Cilliers. Thank-you for taking a chance and opening the door for me, allowing me the opportunity to step into a new world and pursue my passion. I was enthusiastic but unqualified (and ignorant!), yet, over the last three years you have helped me build a solid foundation, gently guiding me, but providing me with the space needed to learn, grow and develop into an independent researcher. I will fondly remember our long meetings and interesting discussions, your

1 < 1% match (Internet from 04-Mar-2019) https://cris.maastrichtuniversity.nl/portal/file/db227ae2-

2 < 1% match (student papers from 31-May-2017) Submitted to University of Cape Town

3 < 1% match (Internet from 30-Aug-2017) http://open.uct.ac.za

4 < 1% match (student papers from 02-May-2018) Submitted to Varsity College

5 < 1% match (publications) Clarissa Wei Shuen Cheong, Elisha Wan Ying Chia, Kuang Teck Tay, Wen Jie Chua et al. "A systematic scoping review of ethical issues in mentoring in internal medicine, family medicine and academic medicine", Advances in Health Sciences Education, 2019

6 < 1% match (Internet from 09-Sep-2019) https://journals.sagepub.com/doi/10.1177/0

7 < 1% match (Internet from 11-Dec-2018) http://iks.ukzn.ac.za

8 < 1% match (publications) Wichard Zwaal, "Hospitality management students' conceptions of education", Tourism and Hospitality Research, 06/2007

References

Adendorff, H. (2011). Strangers in a strange land – on becoming scholars of teaching. *London Review of Education, 9*(3), 305-315. doi:10.1080/14748460.2011.616323

Agarwal, P., Karpicke, J., Kang, S., Roediger, H., & McDermott, K. (2008). Examining the testing effect with open- and closed-book tests. *Applied Cognitive Psychology, 22*(7), 861-876. doi:10.1002/acp.1391

Aker, J., & Mbiti, I. (2010). Mobile Phones and Economic Development in Africa. *Journal of Economic Perspectives, 24*(3), 207-232. doi:10.1257/jep.24.3.207

Åkerlind, G. (2003). Growing and Developing as a University Teacher--Variation in Meaning. *Studies in Higher Education, 28*(4), 375-390. doi:10.1080/0307507032000122242

Åkerlind, G. (2005). Variation and commonality in phenomenographic research methods. *Higher Education Research & Development, 24*(4), 321-334. doi:10.1080/07294360500284672

Åkerlind, G. (2017). What Future for Phenomenographic Research? On Continuity and Development in the Phenomenography and Variation Theory Research Tradition. *Scandinavian Journal of Educational Research, 62*(6), 949-958. doi:10.1080/00313831.2017.1324899

Al Kadri, H., Al-Moamary, M., & van der Vleuten, C. (2009). Students' and teachers' perceptions of clinical assessment program: A qualitative study in a PBL curriculum. *BMC Research Notes, 2*, 263. doi:10.1186/1756-0500-2-263

Albert, M. (2004). Understanding the Debate on Medical Education Research: A Sociological Perspective. *Academic Medicine, 79*(10), 948-954.

Albert, M., Laberge, S., Hodges, B., Regehr, G., & Lingard, L. (2008). Biomedical scientists' perception of the social sciences in health research. *Social Science and Medicine, 66*(12), 2520-2531. doi:10.1016/j.socscimed.2008.01.052

Albert, M., & Reeves, S. (2010). Setting some new standards in medical education research. *Medical Education, 44*(7), 638-639. doi:10.1111/j.1365-2923.2010.03726.x

Alderson, C., & Wall, D. (1992). Does washback exist? *Applied Linguistics, 14*(2), 115-129.

Alexander, E., Osman, N., Walling, J., & Mitchell, V. (2012). Variation and imprecision of clerkship grading in U.S. medical schools. *Academic Medicine, 87*(8), 1070-1076. doi:10.1097/ACM.0b013e31825d0a2a

Anderson, C. (2010). Presenting and Evaluating Qualitative Research. *American Journal of Pharmaceutical Education, 74*(8).

Andrew, N., Ferguson, D., Wilkie, G., Corcoran, T., & Simpson, L. (2009). Developing professional identity in nursing academics: The role of communities of practice. *Nurse Education Today, 29*(6), 607-611.

Archer, R., Elder, W., Hustedde, C., Milam, A., & Joyce, J. (2008). The theory of planned behaviour in medical education: a model for integrating professionalism training. *Medical Education, 42*, 771–777.

Arcila, F. (2018). The Wisdom of Teachers' Personal Theories: Creative ELT Practices From Colombian Rural Schools. *Profile: Issues in Teachers´ Professional Development, 20*(2), 65-78. doi:10.15446/profile.v20n2.67142

Aronson, J. (1994). A Pragmatic View of Thematic Analysis. *The Qualitative Report, 2*(1).

Ashworth, P., & Lucas, U. (1998). What is the 'World' of Phenomenography? *Scandinavian Journal of Educational Research, 42*(4), 415-431. doi:10.1080/0031383980420407

Attia, M., & Edge, J. (2017). Be(com)ing a reflexive researcher: a developmental approach to research methodology. *Open Review of Educational Research, 4*(1), 33-45. doi:10.1080/23265507.2017.1300068

Bahar-Ozvaris, S., Aslan, D., Sahin-Hodoglugil, N., & Sayek, I. (2004). A faculty development program evaluation: from needs assessment to long-term effects, of the teaching skills improvement program. *Teaching and Learning in Medicine, 16*(4), 368-375. doi:10.1207/s15328015tlm1604_11

Bandura, A. (1986). Social cognitive theory. *Annals of child development*.

Barrett, A., Galvin, R., Scherpbier, A., Teunissen, P., O'Shaughnessy, A., & Horgan, M. (2016). Is the learning value of workplace-based assessment being realised? A qualitative study of trainer and trainee perceptions and experiences. *Postgraduate Medical Journal, 93*(1097), 138-142.

Bartholomew, L., Parcel, G., Kok, G., Gottlieb, N., & Fernandez, M. (2011). Planning Health Promotion Programs: An Intervention Mapping Approach. *Wiley*.

Bartle, E., & Thistlethwaite, J. (2014). Becoming a medical educator: motivation, socialisation and navigation. *Medical Education, 14*.

Bearman, M., Dawson, P., Bennett, S., Hall, M., Molloy, E., Boud, D., & Joughin, G. (2017). How university teachers design assessments: a cross-disciplinary study. *Higher Education, 74*(1), 49-64. doi:10.1007/s10734-016-0027-7

Bearman, M., Dawson, P., Boud, D., Bennett, S., Hall, M., & Molloy, E. (2016). Support for assessment practice: developing the Assessment Design Decisions Framework. *Teaching in Higher Education, 21*(5), 545-556. doi:10.1080/13562517.2016.1160217

Beauchamp, C., & Thomas, L. (2009). Understanding teacher identity: an overview of issues in the literature and implications for teacher education. *Cambridge Journal of Education, 39*(2).

Beck, R., Skinner, W., & Schwabrow, L. (2013). A study of sustainable assessment theory in higher education tutorials. *Assessment & Evaluation in Higher Education, 38*(3), 326-348. doi:10.1080/02602938.2011.630978

Beigel, F., Gallardo, O., & Bekerman, F. (2018). Institutional Expansion and Scientific Development in the Periphery: The Structural Heterogeneity of Argentina's Academic Field. *Minerva*. doi:10.1007/s11024-017-9340-2

Berendonk, C., Stalmeijer, R., & Schuwirth, L. (2013). Expertise in performance assessment: assessors' perspectives. *Advances in Health Sciences Education, Theory and Practice, 18*(4), 559-571. doi:10.1007/s10459-012-9392-x

Berry, C., & Taylor, J. (2014). Internationalisation in higher education in Latin America: policies and practice in Colombia and Mexico. *Higher Education, 67*, 585–601.

Bienstock, J., Katz, N., Cox, S., Hueppchen, N., Erickson, S., & Puscheck, E. (2007). To the point: medical education reviews--providing feedback. *American Journal of Obstetrics and Gynecology, 196*(6), 508-513. doi:10.1016/j.ajog.2006.08.021

Bishop, E., & Shepherd, M. (2011). Ethical Reflections: Examining Reflexivity Through the Narrative Paradigm. *Qualitative Health Research, 21*(9), 1283-1294.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). Assessment for learning: putting it into practice. *Open University Press*.

Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7-74. doi:10.1080/0969595980050102

Blitz, J., Bezuidenhout, J., Conradie, H., de Villiers, M., & van Schalkwyk, S. (2014). 'I felt colonised': emerging clinical teachers on a new rural teaching platform. *Rural & Remote Health, 14*(2), 1-9.

Boet, S., Sharma, S., Goldman, J., & Reeves, S. (2012). Review article: Medical education research: an overview of methods. *Journal of Cancer Anesthetics 59*, 159-170.

Bok, H., Teunissen, P., Favier, R., Rietbroek, N., Theyse, L., Brommer, H., . . . Jaarsma, D. (2013). Programmatic assessment of competency-based workplace learning: when theory meets practice. *Medical Education, 13*(1).

Bolander, K., Josephson, A., Mann, S., & Lonka, K. (2006). Teachers Promoting Expertise in Medical Education: understanding the role of the core curriculum. *Quality in Higher Education, 12*(1), 41-55. doi:10.1080/13538320600685156

Bordage, G. (2009). Conceptual frameworks to illuminate and magnify. *Medical Education, 43*(4), 312-319. doi:10.1111/j.1365-2923.2009.03295.x

Bordage, G., Meguerditchian, A., & Tamblyn, R. (2013). Practice indicators of suboptimal care and avoidable adverse events: a content analysis of a national qualifying examination. *Academic Medicine, 88*(10), 1493-1498. doi:10.1097/ACM.0b013e3182a356af

Bottomley, S., & Denny, P. (2011). A participatory learning approach to biochemistry using student authored and evaluated multiple-choice questions. *Biochemical and Molecular Biology Education, 39*(5), 352-361. doi:10.1002/bmb.20526

Boud, D. (1995). Assessment and learning: contradictory or complementary? *Assessment for Learning in Higher Education*, 35-48.

Boud, D. (2000). Sustainable Assessment: Rethinking assessment for the learning society. *Studies in Continuing Education, 22*(2), 151-167. doi:10.1080/713695728

Bowden, J., & Walsh, E. (2000). Phenomenography. *Qualitative Research Methods*.

Bowen, G. (2008). Naturalistic inquiry and the saturation concept: a research note. *Qualitative Research, 8*(1), 137-152. doi:10.1177/1468794107085301

Box, C., Skoog, G., & Dabbs, J. (2015). A Case Study of Teacher Personal Practice Assessment Theories and Complexities of Implementing Formative Assessment. *American Educational Research Journal, 52*(5), 956-983. doi:10.3102/0002831215587754

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101. doi:10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2012). Thematic analysis. In *APA handbook of research methods in psychology, Vol 2: Research designs: Quantitative, qualitative, neuropsychological, and biological.* (pp. 57-71).

Britten, N. (2005). Making sense of qualitative research: a new series. *Medical Education, 39*(1), 5-6. doi:10.1111/j.1365-2929.2004.02024.x

Broadbent, J., Panadero, E., & Boud, D. (2017). Implementing summative assessment with a formative flavour: a case study in a large class. *Assessment and Evaluation in Higher Education*, 1-16. doi:10.1080/02602938.2017.1343455

Broekkamp, H., & van Hout-Wolters, B. (2007). Students' Adaptation of Study Strategies When Preparing for Classroom Tests. *Educational Psychology Review, 19*(4), 401-428. doi:10.1007/s10648-006-9025-0

Brown, G. (2004). Teachers' conceptions of assessment - implications for policy and professional development. *Assessment in Education - Principles, Policy and Practice, 11*(3), 301-318. doi:10.1080/0969594042000304609

Bunniss, S., & Kelly, D. (2010). Research paradigms in medical education research. *Medical Education, 44*(4), 358-366. doi:10.1111/j.1365-2923.2009.03611.x

Burchard, K., Rowland-Morin, P., Coe, N., & Garb, J. (1995). A surgery oral examination: interrater agreement and the influence of rater characteristics. *Academic Medicine, 70*(11), 1044-1046.

Butler, A., & Roediger, H. (2007). Testing improves long-term retention in a simulated classroom setting. *European Journal of Cognitive Psychology, 19*(514-527).

Callahan, C., Hojat, M., & Gonnella, J. (2007). Volunteer bias in medical education research: an empirical study of over three decades of longitudinal data. *Medical Education, 41*(8), 746-753. doi:10.1111/j.1365-2923.2007.02803.x

Cantillon, P., D'Eath, M., de Grave, W., & Dornan, T. (2016). How do clinicians become teachers? A communities of practice perspective. *Advances in Health Sciences Education, 21*(5), 991-1008. doi:10.1007/s10459-016-9674-9

Cantillon, P., Dornan, T., & de Grave, W. (2019). Becoming a Clinical Teacher. *Academic Medicine, 94*(10), 1610-1618. doi:10.1097/acm.0000000000002403

Carless, D. (2007). Learning-oriented assessment: conceptual bases and practical implications. *Innovations in Education and Teaching International, 44*(1), 57-66. doi:10.1080/14703290601081332

Chatzisarantis, N., Hagger, M., Wang, C., & Thøgersen-Ntoumani, C. (2009). The Effects of Social Identity and Perceived Autonomy Support on Health Behaviour Within the Theory of Planned Behaviour. *Current Psychology, 28*(1), 55-68. doi:10.1007/s12144-009-9043-4

Chen, P., & Bonner, S. (2016). Teachers' Beliefs About Grading Practices and a Constructivist Approach to Teaching. *Educational Assessment, 22*(1), 18-34. doi:10.1080/10627197.2016.1271703

Cheng, L. (2000). Washback or Backwash: A Review of the Impact of Testing on Teaching and Learning. *Washback in language testing: Research contexts and methods*.

Chesterman, S., Lamanna, C., Kalamatianou, S., & Rosenstock, T. (2017). Collecting development data with mobile phones: Key considerations from a review of the evidence.

Chokkar, J., Brodbeck, F., & House, R. (2007). Culture and Leadership, Across the World: The GLOBE BOOK of In-depth Studies of 25 Societies. *LEA's Organization and Management Series*.

Cibangu, S., & Hepworth, M. (2016). The uses of phenomenology and phenomenography: A critical review. *Library & Information Science Research, 38*(2), 148-160. doi:10.1016/j.lisr.2016.05.001

Cilliers, F., Schuwirth, L., Adendorff, H., Herman, N., & van der Vleuten, C. (2010). The mechanism of impact of summative assessment on medical students' learning. *Advances in Health Science Education, Theory, and Practice, 15*(5), 695-715. doi:10.1007/s10459-010-9232-9

Cilliers, F., Schuwirth, L., Herman, N., Adendorff, H., & van der Vleuten, C. (2012). A model of the pre-assessment learning effects of summative assessment in medical education. *Advances in Health Science Education, Theory, and Practice, 17*(1), 39-53. doi:10.1007/s10459-011-9292-5

Cilliers, F., Schuwirth, L., & van der Vleuten, C. (2012a). A model of the pre-assessment learning effects of assessment is operational in an undergraduate clinical context. *Medical Education, 12*, 9. doi:10.1186/1472-6920-12-9

Cilliers, F., Schuwirth, L., & van der Vleuten, C. (2012b). Modelling the pre-assessment learning effects of assessment: evidence in the validity chain. *Medical Education, 46*(11), 1087-1098. doi:10.1111/j.1365-2923.2012.04334.x

Cilliers, F., Schuwirth, L., & van der Vleuten, V. (2015). Health behaviour theories: a conceptual lens to explore behaviour change. *Researching Medical Education* 141-153.

Cilliers, F., & Tekian, A. (2016). Effective Faculty Development in an Institutional Context - Designing for Transfer. *Journal of Graduate Medical Education, 8*(2), 145-149. doi:0.4300/JGME-D-15-00117.1

Clarke, V., & Braun, V. (2013). Teaching thematic analysis: Overcoming
challenges and developing strategies for effective learning. *The Psychologist, 26*(2).

Coderre, S., Woloschuk, W., & McLaughlin, K. (2009). Twelve tips for blueprinting. *Medical Teacher, 31*(4), 322-324. doi:10.1080/01421590802225770

Cohen, L., Manion, L., & Morrison, K. (2007). Research Methods in Education.

Comaroff, J., & Comaroff, J. (2012). Theory from the South: Or, how Euro-America is Evolving Toward Africa. *Anthropological Forum, 22*(2), 113-131. doi:10.1080/00664677.2012.694169

Connell, R. (2014). Using southern theory: Decolonizing social thought in theory, research and application. *Planning Theory, 13*(2), 210-223. doi:10.1177/1473095213499216

Connell, R., Collyer, F., Maia, J., & Morrell, R. (2017). Toward a global sociology of knowledge: Post-colonial realities and intellectual practices. *International Sociology, 32*(1), 21-37. doi:10.1177/0268580916676913

Connell, R., Pearse, R., Collyer, F., Maia, J., & Morrell, R. (2018a). Negotiating with the North: How Southern-tier intellectual workers deal with the global economy of knowledge. *The Sociological Review, 66*(1), 41-57.

Connell, R., Pearse, R., Collyer, F., Maia, J., & Morrell, R. (2018b). Re-making the global economy of knowledge: do new fields of research change the structure of North-South relations? *British Journal of Sociology*. doi:10.1111/1468-4446.12294

Cook, D., Bordage, G., & Schmidt, H. (2008). Description, justification and clarification - A framework for classifying the purposes of research in medical education. *Medical Education, 42*(2), 128-133. doi:10.1111/j.1365-2923.2007.02974.x

Couper, M. (2005). Technology Trends in Survey Data Collection. *Social Science Computer Review, 23*(4), 486-501. doi:10.1177/0894439305278972

Creswell, J. (1998). Qualitative inquiry and research design: Choosing among five traditions. *SAGE Publications*.

Creswell, J. (2007). Qualitative Inquiry & Research Design: Choosing Among Five Approaches. *SAGE Publications*.

Creswell, J. (2009). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. *SAGE Publications*.

Crooks, T. (2011). Assessment for learning in the accountability era: New Zealand. *Studies in Educational Evaluation, 37*(1), 71-77. doi:10.1016/j.stueduc.2011.03.002

Crossley, J., Groves, J., Croke, D., & Brennan, P. (2019). Examiner training: A study of examiners making sense of norm-referenced feedback. *Medical Teacher*, 1-8. doi:10.1080/0142159x.2019.1579902

Cruess, R., Cruess, S., Boudreau, J., Snell, L., & Steinert, Y. (2014). Reframing medical education to support professional identity formation. *Academic Medicine, 89*(11), 1446-1451. doi:10.1097/ACM.0000000000000427

Cruess, S., Cruess, R., & Steinert, Y. (2008). Role modelling--making the most of a powerful teaching strategy. *British Medical Journal, 336*(7646), 718-721. doi:10.1136/bmj.39503.757847.BE

Daelmans, H., van der Hem-Stokroos, H., Hoogenboom, R., Scherpbier, A., Stehouwer, C., & van der Vleuten, C. (2005). Global clinical performance rating, reliability and validity in an undergraduate clerkship. *Netherlands Journal of Medicine, 63*(7), 279-284.

Darawsheh, W. (2014). Reflexivity in research: Promoting rigour, reliability and validity in qualitative research. *International Journal of Therapy and Rehabilitation, 21*(12), 560-568.

Davis, D. (2018). The medical school without walls: Reflections on the future of medical education. *Medical Teacher*, 1-6. doi:10.1080/0142159X.2018.1507263

Davis, R., Campbell, R., Hildon, Z., Hobbs, L., & Michie, S. (2015). Theories of behaviour and behaviour change across the social and behavioural sciences: a scoping review. *Health Psychology Review, 9*(3), 323-344. doi:10.1080/17437199.2014.941722

Dawson, P., Bearman, M., Boud, D., Hall, M., Molloy, E., Bennett, S., & Joughin, G. (2013). Assessment might dictate the curriculum, but what dictates assessment? *Teaching and Learning Inquiry - The ISSOTL Journal, 1*(1), 107-111.

de Jonge, L., Timmerman, A., Govaerts, M., Muris, J., Muijtjens, A., Kramer, A., & van der Vleuten, C. (2017). Stakeholder perspectives on workplace-based performance assessment: towards a better understanding of assessor behaviour. *Advances in Health Science Education, Theory, and Practice*. doi:10.1007/s10459-017-9760-7

DeLuca, C., Coombs, A., & LaPointe-McEwan, D. (2019). Assessment mindset: Exploring the relationship between teacher mindset and approaches to classroom assessment. *Studies in Educational Evaluation, 61*, 159-169. doi:10.1016/j.stueduc.2019.03.012

Deneen, C., & Boud, D. (2013). Patterns of resistance in managing assessment change. *Assessment & Evaluation in Higher Education, 39*(5), 577-591. doi:10.1080/02602938.2013.859654

Denscombe, M. (2010). The good research guide for small-scale social research projects.

Dicicco-Bloom, B., & Crabtree, B. (2006). The qualitative research interview. *Medical Education, 40*(4), 314-321. doi:10.1111/j.1365-2929.2006.02418.x

Dochy, F., Segers, M., Gijbels, D., & Struyven, K. (2007). Assessment Engineering: Breaking down barriers between teaching and learning, and assessment. *Rethinking assessment in higher education: learning for the longer term*, 87-100.

Doja, A., Horsley, T., & Sampson, M. (2014). Productivity in medical education research: an examination of countries of origin. *Medical Education, 14*, 243.

Dornan, T., Peile, E., & Spencer, J. (2009). In defence of the existing strengths of medical education research. *Medical Education, 43*(4), 391. doi:10.1111/j.1365-2923.2009.03305.x

Dornan, T., & Spencer, E. (2008). On 'evidence'. *Medical Education, 42*(3), 232-234. doi:10.1111/j.1365-2923.2007.02997.x

Dortins, E. (2002). Reflections on phenomenographic process: Interview, transcription and analysis. *Higher Education Research and Development Society of Australasia*, 207-213.

Downing, S. (2003). Validity: on the meaningful interpretation of assessment data. *Medical Education, 37*(9), 830-837.

Downing, S. (2004). Reliability: on the reproducibility of assessment data. *Medical Education, 38*(9), 1006-1012. doi:10.1046/j.1365-2929.2004.01932.x

Downing, S., & Yudkowsky, R. (2009). Introduction to assessment in the health professions. *Assessment in Health Professions Education*, 1-21.

Dreyfus, S. (2004). The Five-Stage Model of Adult Skill Acquisition. *Bulletin of Science, Technology & Society, 24*(3), 177-181. doi:10.1177/0270467604264992

Driessen, E., van Tartwijk, J., Govaerts, M., Teunissen, P., & van der Vleuten, C. (2012). The use of programmatic assessment in the clinical workplace: a Maastricht case report. *Medical Teacher, 34*(3), 226-231. doi:10.3109/0142159X.2012.652242

Earley, P., & Stubblebine, P. (1989). Intercultural Assessment of Performance Feedback. *Group & Organization management*.

Eaton, L., Flisher, A., & Aarø, L. (2003). Unsafe sexual behaviour in South African youth. *Social Science & Medicine, 56*, 149-165.

Ekeblad, E. (1997). On the surface of phenomenography: A response to Graham Webb. *Higher Education, 33*, 219-224.

Engward, H., & Davis, G. (2015). Being reflexive in qualitative grounded theory: discussion and application of a model of reflexivity. *Journal of Advanced Nursing, 71*(7), 1530-1538. doi:10.1111/jan.12653

Entwistle, N. (1997a). Introduction: Phenomenography in Higher Education. *Higher Education Research & Development*, 127-134. doi:10.1080/0729436970160202

Entwistle, N. (1997b). Reconstituting approaches to learning: A response to Webb. *Higher Education, 33*, 213-218.

Entwistle, N., & Peterson, E. (2004). Conceptions of learning and knowledge in higher education: Relationships with study behaviour and influences of learning environments. *International Journal of Educational Research, 41*(6), 407-428. doi:10.1016/j.ijer.2005.08.009

Ericsson, A. (1998). The Scientific Study of Expert Levels of Performance: general implications for optimal learning and creativity. *High Ability Studies, 9*(1), 75-100. doi:10.1080/1359813980090106

Eva, K. (2009). Broadening the debate about quality in medical education research. *Medical Education, 43*(4), 294-296. doi:10.1111/j.1365-2923.2009.03342.x

Fawcett, B., & Hearn, J. (2004). Researching others: Epistemology, experience, standpoints and participation. *International Journal of Social Research Methodology: Theory & Practice, 7*(3), 201-218. doi:0.1080/13645570210163989

Fereday, J. (2006). Demonstrating Rigor Using Thematic Analysis: A Hybrid Approach of Inductive and Deductive Coding and Theme Development. *International Journal of Qualitative Methods, 5*(1), 80-92.

Fielding, K., McDonald, R., & Louis, W. (2008). Theory of planned behaviour, identity and intentions to engage in environmental activism. *Journal of Environmental Psychology, 28*(4), 318-326.

Finlay, L. (2002a). Negotiating the swamp: the opportunity and challenge of reflexivity in research practice. *Qualitative Research, 2*(2), 209-230.

Finlay, L. (2002b). "Outing" the Researcher: The Provenance, Process, and Practice of Reflexivity. *Qualitative Health Research, 12*, 4.

Fletcher, R., Meyer, L., Anderson, H., Johnston, P., & Rees, M. (2011). Faculty and Students Conceptions of Assessment in Higher Education. *Higher Education, 64*(1), 119-133. doi:10.1007/s10734-011-9484-1

Foley, D. (2003). Indigenous Epistemolgy and Indigenous Standpoint Theory. *Social Alternative, 22*(1), 44-54.

Foley, D. (2006). Indigenous Standpoint Theory. *International Journal of the Humanities, 3*(8), 25-26.

Gandomkar, R., & Sandars, J. (2017). The importance of context in medical education program evaluation. *Medical Teacher*, 1-1. doi:10.1080/0142159x.2017.1390215

Gibbs, G., & Dunbar-Goddet, H. (2009). Characterising programme-level assessment environments that support learning. *Assessment & Evaluation in Higher Education, 34*(4), 481-489. doi:10.1080/02602930802071114

Gibbs, G., & Simpson, C. (2004). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education, 1*(1), 3-31.

Gibbs, T., Durning, S., & van der Vleuten, C. (2011). Theories in medical education: towards creating a union between educational practice and research traditions. *Medical Teacher, 33*(3), 183-187. doi:10.3109/0142159X.2011.551680

Gilles, J., Detroz, P., & Blais, J. (2011). An international online survey of the practices and perceptions of higher education professors with respect to the assessment of learning in the classroom. *Assessment and Evaluation in Higher Education, 36*(6), 719-733. doi:10.1080/02602938.2010.484880

Gillsjö, C., Schwartz-Barcott, D., Bergh, I., & Dahlgren, L. (2011). Older Adults' Ways of Dealing With Daily Life While Living With Long-Term Musculoskeletal Pain at Home. *Journal of Applied Gerontology, 31*(5), 685-705. doi:10.1177/0733464810397540

Glanz, K., & Bishop, D. (2010). The role of behavioral science theory in development and implementation of public health interventions. *Annual Review of Public Health, 31*, 399-418. doi:10.1146/annurev.publhealth.012809.103604

Glanz, K., Rimer, B., & Viswanath, K. (2008). Health Behavior and Health Education. *Jossey-Bass, San Francisco*.

Godin, G., Belanger-Gravel, A., Eccles, M., & Grimshaw, J. (2008). Healthcare professionals' intentions and behaviours: a systematic review of studies based on social cognitive theories. *Implementation Sciences, 3*, 36. doi:10.1186/1748-5908-3-36

Goos, M., & Hughes, C. (2010). An investigation of the confidence levels of course/subject coordinators in undertaking aspects of their assessment responsibilities. *Assessment and Evaluation in Higher Education, 35*(3), 315-324. doi:10.1080/02602930903221477

Gosselin, K., Norris, J., & Ho, M. (2016). Beyond homogenization discourse: Reconsidering the cultural consequences of globalized medical education. *Medical Teacher, 38*(7), 691-699. doi:10.3109/0142159X.2015.1105941

Govaerts, M., van de Wiel, M., Schuwirth, L., van der Vleuten, C., & Muijtjens, A. (2013). Workplace-based assessment: raters' performance theories and constructs. *Advances in Health Sciences Education, Theory and Practice, 18*(3), 375-396. doi:10.1007/s10459-012-9376-x

Greysen, S., Dovlo, D., Olapade-Olaopa, E., Jacobs, M., Sewankambo, N., & Mullan, F. (2011). Medical education in sub-Saharan Africa: a literature review. *Medical Education, 45*(10), 973-986. doi:10.1111/j.1365-2923.2011.04039.x

Grossman, R., & Salas, E. (2011). The transfer of training - what really matters. *International Journal of Training and Development, 15*(2), 103-120.

GSMA. (2016). The Mobile Economy Africa. . *GSMA Intelligence, London, United Kingdom.* .

Guba, E., & Lincoln, Y. (1981). Effective evaluation: Improving the usefulness of evaluation results through responsive and naturalistic approaches. *he Jossey-Bass higher and adult education series and the Jossey-Bass social and behavioral science series*.

Guba, E., & Lincoln, Y. (1982). Epistemological and methodological bases of naturalistic inquiry. *Education and Communication Technology: A Journal of Theory, Research, and Development, 30*(4), 233–252.

Guest, G., Bunce, A., & Johnson, L. (2016). How Many Interviews Are Enough? *Field Methods, 18*(1), 59-82. doi:10.1177/1525822x05279903

Guest, G., MacQueen, K., & Namey, E. (2012). *Applied Thematic Analysis*.

Guillemin, M., & Gillam, L. (2016). Ethics, Reflexivity, and "Ethically Important Moments" in Research. *Qualitative Inquiry, 10*(2), 261-280. doi:10.1177/1077800403262360

Hagger, M., & Chatzisarantis, N. (2006). Self-identity and the theory of planned behaviour: Between- and within-participants analyses. *British Journal of Social Psychology, 45*(4), 731-757. doi:10.1348/014466605x85654

Halinen, K., Ruohoniemi, M., Katajavuori, N., & Virtanen, V. (2013). Life science teachers' discourse on assessment: a valuable insight into the variable conceptions of assessment in higher education. *Journal of Biological Education, 48*(1), 16-22. doi:10.1080/00219266.2013.799082

Hamdy, H. (2006). Blueprinting for the assessment of health care professionals. *The Clinical Teacher, 3*, 175-179.

Hanesworth, P., Bracken, S., & Elkington, S. (2018). A typology for a social justice approach to assessment: learning from universal design and culturally sustaining pedagogy. *Teaching in Higher Education*, 1-17. doi:10.1080/13562517.2018.1465405

Harden, R., & Lilley, P. (2018). Best practice for assessment. *Medical Teacher*, 1-3. doi:10.1080/0142159x.2018.1513804

Harding, S. (2009). Standpoint Theories: Productively Controversial. *Hypatia, 24*(4), 192-200.

Harland, T., McLean, A., Wass, R., Miller, E., & Sim, K. (2014). An assessment arms race and its fallout: high-stakes grading and the case for slow scholarship. *Assessment & Evaluation in Higher Education, 40*(4), 528-541. doi:10.1080/02602938.2014.931927

Harris, L. (2011). Phenomenographic perspectives on the structure of conceptions: The origins, purposes, strengths, and limitations of the what/how and referential/structural frameworks. *Educational Research Review, 6*(2), 109-124. doi:10.1016/j.edurev.2011.01.002

Harrison, C., Konings, K., Schuwirth, L., Wass, V., & van der Vleuten, C. (2017). Changing the culture of assessment: the dominance of the summative assessment paradigm. *Medical Education, 17*(1), 73. doi:10.1186/s12909-017-0912-5

Hartman, S., & Nelson, M. (1992). What we say and what we do: self-reported teaching behaviour versus performances in written simulations among medical school faculty. *Academic Medicine, 67*, 522-527.

Heeneman, S., Oudkerk Pool, A., Schuwirth, L., van der Vleuten, C., & Driessen, E. (2015). The impact of programmatic assessment on student learning: theory versus practice. *Medical Education, 49*(5), 487-498. doi:10.1111/medu.12645

Helsing, D., Howell, A., Kegan, R., & Lahey, L. (2008). Putting the "Development" in Professional Development: Understanding and Overturning Educational Leaders' Immunities to Change. *Harvard Educational Review, 78*(3).

Henderson, M., Ryan, T., & Phillips, M. (2019). The challenges of feedback in higher education. *Assessment & Evaluation in Higher Education*, 1-16. doi:10.1080/02602938.2019.1599815

Henrich, J., Heine, S., & Norenzayan, A. (2010). The weirdest people in the world? *RatSWD Working Paper Series 139*.

Ho, A. (2000). A conceptual change approach to staff development A model for programme design. *International Journal for Academic Development, 5*(1), 30-41. doi:10.1080/136014400410088

Ho, A., Watkins, D., & Kelly, M. (2001). The Conceptual Change Approach to Improving Teaching and Learning: An Evaluation of a Hong Kong Staff Development Programme. *Higher Education, 42*(2), 143-169.

Hofstede, G. (2001). Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations Across Nations. Second Edition. *Thousand Oaks CA: Sage Publications*.

Hofstede, G. (2006). What Did GLOBE Really Measure? Researchers' Minds versus Respondents' Minds. *Journal of International Business Studies, 37*(6), 882-896.

Hofstede, G. (2010). The GLOBE debate- Back to relevance. *Journal of International Business Studies, 41*(8), 1339-1346.

Hofstede, G. (2011). Dimensionalizing Cultures: The Hofstede Model in Context. *Online Readings in Psychology and Culture, 2*(1). doi:10.9707/2307-0919.1014

Hofstede, G., & McCrae, R. (2004). Personality and Culture Revisited: Linking Traits and Dimensions of Culture. *Cross-Cultural Research, 38*(1), 52-88. doi:10.1177/1069397103259443

Hofstede, G., & Minkov, M. (2010). Cultures and organizations: Software of the mind. Revised and expanded. . *McGraw-Hill, New York.*

Hofstede, G., Neuijen, B., Ohayv, D., & Sanders, D. (1990). Measuring Organizational Cultures: A Qualitative and Quantitative Study Across Twenty Cases. *Administrative Science Quarterly, 35*(2), 286-316.

Holland, K., Middleton, L., & Uys, L. (2013). Professional Confidence: Conceptions Held by Novice Occupational Therapists in South Africa. *Occupational Therapy International, 20*, 105-113.

Holmboe, E. (2004). Faculty and the observation of trainees' clinical skills- problems and opportunities. *Academic Medicine, 79*(1), 16-22.

Holmboe, E., Ward, D., Reznick, R., Katsufrakis, P., Leslie, K., Patel, V., . . . Nelson, E. (2011). Faculty development in assessment: the missing link in competency-based medical education. *Academic Medicine, 86*(4), 460-467. doi:10.1097/ACM.0b013e31820cb2a7

Homer, M., Fuller, R., & Pell, G. (2017). The benefits of sequential testing: Improved diagnostic accuracy and better outcomes for failing students. *Medical Teacher*, 1-10. doi:10.1080/0142159X.2017.1404561

Hopkins, R., Regehr, G., & Pratt, D. (2017). A framework for negotiating positionality in phenomenological research. *Medical Teacher, 39*(1), 20-25. doi:10.1080/0142159X.2017.1245854

House, R., Dorfman, P., Javidan, M., Hanges, P., & de Luque, M. (2014). Strategic Leadership Across Cultures: The GLOBE Study of CEO Leadership Behavior and Effectiveness in 24 Countries 2014.

House, R., Hanges, P., Javidan, M., Dorfman, P., & Gupta, V. (2004). Culture, Leadership, and Organizations: The GLOBE Study of 62 Societies.

Hunt, S. (1989). Naturalistic, Humanistic, and Interpretive Inquiry: Challenges and Ultimate Potential. *Interpretive Consumer Research*, 185-198.

Huwendiek, S., Mennin, S., Dern, P., Ben-David, M., van der Vleuten, C., Tonshoff, B., & Nikendei, C. (2010). Expertise, needs and challenges of medical educators: Results of an international web survey. *Medical Teacher, 32*(11), 912-918. doi:10.3109/0142159X.2010.497822

Isaacs, T., Zara, C., Herbert, G., Coombs, S., & Smith, C. (2013). Key Concepts in Educational Assessment. *SAGE Publications*.

ITU. (2016). Data: Mobile-cellular telephone subscriptions 2000-2015. .

James, N., & Busher, H. (2016). Qualitative Research Methods. Chapter 15: Online Interviewing. *Qualitative Research Methods*.

Janghorban, R., Latifnejad Roudsari, R., & Taghipour, A. (2014). Skype interviewing: the new generation of online synchronous interview in qualitative research. *International Journal of Qualitative Studies on Health and Well-being, 9*, 24152. doi:10.3402/qhw.v9.24152

Javidan, M., House, R., Dorfman, P., Hanges, P., & de Luque, M. (2006). Conceptualizing and Measuring Cultures and Their Consequences: A Comparative Review of GLOBE's and Hofstede's Approaches. *Journal of International Business Studies, 37*(6), 897-914.

Jenner, B., & Myers, K. (2018). Intimacy, rapport, and exceptional disclosure: a comparison of in-person and mediated interview contexts. *International Journal of Social Research Methodology, 22*(2), 165-177. doi:10.1080/13645579.2018.1512694

Jessop, T., El Hakim, Y., & Gibbs, G. (2013). The whole is greater than the sum of its parts: a large-scale study of students' learning in response to different programme assessment

patterns. *Assessment & Evaluation in Higher Education, 39*(1), 73-88. doi:10.1080/02602938.2013.792108

Jessop, T., & Tomas, C. (2016). The implications of programme assessment patterns for student learning. *Assessment and Evaluation in Higher Education, 42*(6), 990-999. doi:10.1080/02602938.2016.1217501

Johnson, G., Scholes, K., & Whittington, R. (2005). Exploring Corporate Strategy: Text and Cases.

Jootun, D., McGhee, G., & Marland, G. (2009). Reflexivity: promoting rigour in qualitative research. *Nursing Standard, 23*(23), 42-46.

Jowett, A., Peel, E., & Shaw, R. (2011). Online Interviewing in Psychology: Reflections on the Process. *Qualitative Research in Psychology, 8*(4), 354-369. doi:10.1080/14780887.2010.500352

Kálmán, O., Tynjälä, P., & Skaniakos, T. (2019). Patterns of university teachers' approaches to teaching, professional development and perceived departmental cultures. *Teaching in Higher Education*, 1-20. doi:10.1080/13562517.2019.1586667

Kane, R., Sandretto, S., & Heath, C. (2002). Telling half the story: A critical review of research on the teaching beliefs and practices of university academics. *Review of Educational Research, 72*(2), 177-122.

Karlsson, G. (2009). Correlating facts or interpreting meaning: Two different epistemological projects within medical research. *Scandinavian Journal of Occupational Therapy*, 68-75.

Karpicke, J., Butler, A., & Roediger, H. (2009). Metacognitive strategies in student learning: do students practise retrieval when they study on their own? *Memory, 17*(4), 471-479. doi:10.1080/09658210802647009

Karpicke, J., & Roediger, H. (2008). The Critical Importance of Retrieval for Learning. *Science, 319*, 966-968.

Karthikeyan, S., O'Connor, E., & Hu, W. (2019). Barriers and facilitators to writing quality items for medical school assessments - a scoping review. *Medical Education, 19*(1), 123. doi:10.1186/s12909-019-1544-8

Kay, D., Berry, A., & Coles, N. (2019). What Experiences in Medical School Trigger Professional Identity Development? *Teaching and Learning in Medicine, 31*(1), 17-25. doi:10.1080/10401334.2018.1444487

Kember, D. (1997). A reconceptualisation of the research into university academics' conceptions of teaching. *Learning and Instruction, 7*(3), 255-275.

Kester, K. (2018). Reproducing peace? A CRT analysis of Whiteness in the curriculum and teaching at a University of The UN. *Teaching in Higher Education, 24*(2), 212-230. doi:10.1080/13562517.2018.1476339

Kluger, A., & DeNisi, A. (1996). The Effects of Feedback Interventions on Performance: A Historical Review, a Meta-Analysis, and a Preliminary Feedback Intervention Theory. *Psychological Bulletin, 119*(2), 254-284.

Knight, P., & Trowler, P. (2000). Department-level Cultures and the Improvement of Learning and Teaching. *Studies in Higher Education, 25*(1), 69-83. doi:0.1080/030750700116028

Knox, S., & Burkard, A. (2009). Qualitative research interviews. *Psychotherapy Research, 19*(4-5), 566-575. doi:10.1080/10503300802702105

Koch, T., & Harrington, A. (1998). Reconceptualizing rigour: the case for reflexivity. *Journal of Advanced Nursing, 28*(4), 882-890.

Kogan, J., Conforti, L., Yamazaki, K., Iobst, W., & Holmboe, E. (2017). Commitment to Change and Challenges to Implementing Changes After Workplace-Based Assessment Rater Training. *Academic Medicine, 92*(3), 394-402.

Kok, G. (2014). A practical guide to effective behaviour change: How to apply theory- and evidence-based behaviour change methods in an intervention. *The European Health Psychologist, 16*(5), 156-170.

Kok, G., Gottlieb, N., Peters, G., Mullen, P., Parcel, G., Ruiter, R., . . . Bartholomew, L. (2016). A taxonomy of behaviour change methods: an Intervention Mapping approach. *Health Psychology Review, 10*(3), 297-312. doi:10.1080/17437199.2015.1077155

Konradt, U., Otte, K., Schippers, M., & Steenfatt, C. (2016). Reflexivity in Teams: A Review and New Perspectives. *Journal of Psychology, 150*(2), 153-174. doi:10.1080/00223980.2015.1050977

Korstjens, I., & Moser, A. (2018). Series: Practical guidance to qualitative research. Part 4: Trustworthiness and publishing. *European Journal of General Practice, 24*(1), 120-124. doi:10.1080/13814788.2017.1375092

Kreutzer, T. (2009). Generation Mobile: Online and Digital Media Usage on Mobile Phones among Low-Income Urban Youth in South Africa.

Kvale, S. (1992). Ten standard responses to Qualitative Research Interviews. *Journal of Phenomenological Psychology, 14*(1-2), 171-196.

Kvale, S. (2006). Dominance Through Interviews and Dialogues. *Qualitative Inquiry, 12*(3), 480-500.

Larsen, D., Butler, A., & Roediger, H. (2008). Test-enhanced learning in medical education. *Medical Education, 42*(10), 959-966. doi:10.1111/j.1365-2923.2008.03124.x

Larsson, J., & Holmström, I. (2007). Phenomenographic or phenomenological analysis: does it matter? Examples from a study on anaesthesiologists' work. *International Journal of Qualitative Studies on Health and Well-being, 2*(1), 55-64. doi:0.1080/17482620601068105

Larsson, J., Holmström, I., & Rosenqvist, U. (2003). Professional artist, good Samaritan, servant and co-ordinator- four ways of understanding the anaesthetist's work. *Acta Anaesthesiologica Scandinavica, 47*(7), 787-793.

le Roux, K. (2018). Navigating the discursive tensions of the public university: the case of a South African student. *Teaching in Higher Education, 24*(2), 231-247. doi:10.1080/13562517.2018.1485644

Leibowitz, B. (2017). Cognitive justice and the higher education curriculum. *Journal of Education (University of Kwa Zulu Natal), 68*.

Levitt, H., Motulsky, S., Wertz, F., Morrow, S., & Ponterotto, J. (2017). Recommendations for designing and reviewing qualitative research in psychology: Promoting methodological integrity. *Qualitative Psychology, 4*(1), 2-22. doi:10.1037/qup0000082

Lieff, S., Baker, L., Mori, B., Egan-Lee, E., Chin, K., & Reeves, S. (2012). Who am I? Key influences on the formation of academic identity within a faculty development program. *Medical Teacher, 34*(3), e208-215. doi:10.3109/0142159X.2012.642827

Limberg, L. (2000). Phenomenography: a relational approach to research on information needs, seeking and sue. *The New Review of Information Behaviour Research*, 51-67.

Lippke, S., & Ziegelmann, J. (2008). Theory-Based Health Behavior Change: Developing, Testing, and Applying Theories for Evidence-Based Interventions. *Applied Psychology, 57*(4), 698-716. doi:10.1111/j.1464-0597.2008.00339.x

Lo Iacono, V., Symonds, P., & Brown, D. (2016). Skype as a Tool for Qualitative Research Interviews. *Sociological Research Online, 21*(2), 1-15. doi:10.5153/sro.3952

Lord, J., Mourtzanos, E., McLaren, K., Murray, S., Kimmel, R., & Cowley, D. (2012). A peer mentoring group for junior clinician educators: four years' experience. *Academic Medicine, 87*(3), 378-383. doi:10.1097/ACM.0b013e3182441615

Luckett, K. (2016). Curriculum contestation in a post-colonial context: a view from the South. *Teaching in Higher Education, 21*(4), 415-428. doi:10.1080/13562517.2016.1155547

Major, C., & Savin-Baden, M. (2011). Integration of qualitative evidence: towards construction of academic knowledge in social science and professional fields. *Qualitative Research, 11*(6), 645-663. doi:10.1177/1468794111413367

Makary, M., & Daniel, M. (2016). Medical error-the third leading cause of death in the US. *British Medical Journal, 353*, i2139. doi:10.1136/bmj.i2139

Maldonado-Torres, N. (2016). Outline of Ten Theses on Coloniality and Decoloniality. *Foundation: Franz Fanon*, 1-37.

Mälkki, K., & Lindblom-Ylänne, S. (2012). From reflection to action? Barriers and bridges between higher education teachers' thoughts and actions. *Studies in Higher Education, 37*(1), 33-50. doi:10.1080/03075079.2010.492500

Malterud, K. (2001). Qualitative research: standards, challenges, and guidelines. *The Lancet, 358*(9280), 483-488. doi:10.1016/s0140-6736(01)05627-6

Mann, K., MacLeod, A., Cleland, J., & Durning, S. (2015). Chapter 6: Constructivism: learning theories and approaches to research. *Researching Medical Education*, 51-65.

Marton, F. (1981). Phenomenography – describing conceptions of the world around us *Instructional Science 10*, 177-200.

Marton, F., & Pong, W. (2005). On the unit of description in phenomenography. *Higher Education Research & Development, 24*(4), 335-348. doi:10.1080/07294360500284706

Marton, F., & Saljo, R. (1976). On Qualitative Differences in Learning: Outcome and Process. *British Journal of Educational Psychology*. doi:https://doi.org/10.1111/j.2044-8279.1976.tb02304.x

Massie, J., & Ali, J. (2016). Workplace-based assessment: a review of user perceptions and strategies to address the identified shortcomings. *Advances in Health Sciences Education, Theory and Practice, 21*(2), 455-473. doi:10.1007/s10459-015-9614-0

Maxwell, J. (2004a). Causal Explanation, Qualitative Research, and Scientific Inquiry in Education. *Educational Researcher*, 1-11.

Maxwell, J. (2004b). Using Qualitative Methods for Causal Explanation. *Field Methods, 16*(3), 243-264. doi:10.1177/1525822x04266831

McGaghie, W. (2015). Varieties of integrative scholarship: why rules of evidence, criteria, and standards matter. *Acad Med, 90*(3), 294-302. doi:10.1097/ACM.0000000000000585

McGill, D., van der Vleuten, C., & Clarke, M. (2013). A critical evaluation of the validity and the reliability of global competency constructs for supervisor assessment of junior medical trainees. *Advances in Health Science Education, Theory, and Practice, 18*(4), 701-725. doi:10.1007/s10459-012-9410-z

McGrath, C., Palmgren, P., & Liljedahl, M. (2018). Twelve tips for conducting qualitative research interviews. *Medical Teacher*, 1-5. doi:10.1080/0142159x.2018.1497149

McKimm, J. (2009). Giving effective feedback. *British Journal of Hospital Medicine, 70*(3), 158-161.

McKinley, D., & Norcini, J. (2014). How to set standards on performance-based examinations: AMEE Guide No. 85. *Medical Teacher, 36*(2), 97-110. doi:10.3109/0142159X.2013.853119

McLeay, F., & Wesson, D. (2014). Chinese versus UK marketing students' perceptions of peer feedback and peer assessment. *The International Journal of Management Education, 12*(2), 142-150. doi:10.1016/j.ijme.2014.03.005

McLeod, P., Meagher, T., Steinert, Y., Schuwirth, L., & McLeod, A. (2004). The Clinical TeacherClinical teachers' tacit knowledge of basic pedagogic principles. *Medical Teacher, 26*(1), 23-27. doi:10.1080/01421590310001643154

McMillan, W., Cleland, J., & Durning, S. (2015). Chapter 2: Theory in healthcare education research: the importance of worldview. *Researching Medical Education*, 15-23.

Medland, E. (2018). 'I'm an assessment illiterate': towards a shared discourse of assessment literacy for external examiners. *Assessment & Evaluation in Higher Education*, 1-16. doi:10.1080/02602938.2018.1523363

Meijer, M., Kuijpers, M., Boei, F., Vrieling, E., & Geijsel, F. (2016). Professional development of teacher-educators towards transformative learning. *Professional Development in Education, 43*(5), 819-840. doi:10.1080/19415257.2016.1254107

Mezirow, J. (1990). Fostering Critical Reflection in Adulthood: A Guide to Transformative and Emancipatory Learning.

Mezirow, J. (1997). Transformative Learning- Theory to Practice Transformative Learning Theory. . *Transformative Learning, 74*, 5-12.

Michie, S., Johnston, M., Abraham, C., Lawton, R., Parker, D., & Walker, A. (2005). Making psychological theory useful for implementing evidence based practice: a consensus approach. *Quality and Safety in Health Care, 14*(1), 26-33. doi:10.1136/qshc.2004.011155

Miles, M., & Huberman, A. (1994). An expanded sourcebook: Qualitative Data Analysis. *SAGE Publications*.

Moller, M., Fridlund, B., & Goransson, K. (2010). Patients' conceptions of the triage encounter at the Emergency Department. *Scandinavian Journal of Caring Science, 24*(4), 746-754. doi:10.1111/j.1471-6712.2010.00772.x

Monrouxe, L. (2010). Identity, identification and medical education: why should we care? *Medical Education, 44*, 40-49.

Monrouxe, L., & Rees, C. (2009). Picking up the gauntlet: constructing medical education as a social science. *Medical Education, 43*(3), 196-198. doi:10.1111/j.1365-2923.2008.03272.x

Montgomery, C. (2019). Surfacing "Southern" Perspectives on Student Engagement With Internationalization: Doctoral Theses as Alternative Forms of Knowledge. *Journal of Studies in International Education, 23*(1), 123-138.

Morrell, R. (2016). Making Southern theory? Gender researchers in South Africa. *Feminist Theory, 17*(2), 191-209. doi:10.1177/1464700116645877

Morse, J. (1991). Strategies for sampling. *Qualitative nursing research: A contemporary dialogue*, 127-145.

Morse, J. (1994). “Emerging from the data”: The cognitive process of analysis in qualitative inquiry. *Critical issues in qualitative research methods*, 23-46.
Morse, J. (1995). The significance of saturation. *Qualitative Health Research, 5*(2), 147-149.
Morse, J. (2010). "Cherry picking": writing from thin data. *Qualitative Health Research, 20*(1), 3. doi:10.1177/1049732309354285
Morse, J. (2015a). Critical Analysis of Strategies for Determining Rigor in Qualitative Inquiry. *Qualitative Health Research, 25*(9), 1212-1222. doi:10.1177/1049732315588501
Morse, J. (2015b). “Data Were Saturated . . . ”. *Qualitative Health Research, 25*(5), 587-588.
Morse, J., Barrett, R., Mayan, M., Olson, K., & Spiers, J. (2002). Verification Strategies for Establishing Reliability and Validity in Qualitative Research. *International Journal of Qualitative Methods, 1*(2), 13-22.
Moss, P., Phillips, D., Erickson, F., Floden, R., Lather, P., & Schneider, B. (2009). Learning From Our Differences: A Dialogue Across Perspectives on Quality in Education Research. *Educational Researcher, 38*(7), 501-517. doi:10.3102/0013189x09348351
Mruck, K., & Breuer, F. (2003). Subjectivity and Reflexivity in Qualitative Research—The FQS Issues. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research, 4*(2).
Mullan, F., Frehywot, S., Omaswa, F., Buch, E., Chen, C., Greysen, R., . . . Neusy, A. (2011). Medical schools in sub-Saharan Africa. *The Lancet, 377*(9771), 1113-1121. doi:10.1016/s0140-6736(10)61961-7
Munck, B., Sandgren, A., Fridlund, B., & Martensson, J. (2012a). Next-of-kin's conceptions of medical technology in palliative homecare. *Journal of Clinical Nursing, 21*(13-14), 1868-1877. doi:10.1111/j.1365-2702.2012.04123.x
Munck, B., Sandgren, A., Fridlund, B., & Martensson, J. (2012b). Patients’ Understanding of Medical Technology in Palliative Home Care. *Journal of Hospice & Palliative Nursing, 14*(3), 191-198. doi:10.1097/NJH.0b013e3182404b71
Munro, S., Lewin, S., Swart, T., & Volmink, J. (2007). A review of health behaviour theories: how useful are these for developing interventions to promote long-term medication adherence for TB and HIV/AIDS? *Public Health, 7*(1). doi:10.1186/1471-2458-7-104
Murray, K., & Macdonald, R. (1997). The disjunction between lecturers' conceptions of teaching and their claimed educational practice. *Higher Education, 33*, 331-349.
Myyry, L., Karaharju-Suvanto, T., Vesalainen, M., Virtala, A., Raekallio, M., Salminen, O., . . . Nevgi, A. (2019). Experienced academics' emotions related to assessment. *Assessment & Evaluation in Higher Education*, 1-13. doi:10.1080/02602938.2019.1601158
Natriello, G. (1987). The impact of evaluation processes on students. *Educational Psychologist, 22*(2), 155-175.
Ndlovu-Gatsheni, S. (2013). Why decoloniality in the 21st century? *The Thinker, 48*(10), 10-15.
Newble, D. (2016). Revisiting 'The effect of assessments and examinations on the learning of medical students'. *Medical Education, 50*(5), 498-501. doi:10.1111/medu.12796
Ng, K., Koh, C., Ang, S., Kennedy, J., & Chan, K. (2011). Rating Leniency and Halo in Multisource Feedback Ratings: Testing Cultural Assumptions of Power Distance and Individualism-Collectivism. *Journal of Applied Psychology, 96*(5), 1033–1044.
Niyibizi, E., Sibomana, E., Niyomugabo, C., Yanzigiye, B., Jean de Dieu, A., & Perumal, J. (2018). Assessment in a Rwandan higher education institution: a quest for aligned assessment to

promote socio-economic transformation. *Assessment & Evaluation in Higher Education, 43*(7), 1166-1182. doi:10.1080/02602938.2018.1436688

Norcini, J., Anderson, B., Bollela, V., Burch, V., Costa, M., Duvivier, R., . . . Roberts, T. (2011). Criteria for good assessment: consensus statement and recommendations from the Ottawa 2010 Conference. *Medical Teacher, 33*(3), 206-214. doi:10.3109/0142159X.2011.551559

Norton, L., Floyd, S., & Norton, B. (2019). Lecturers' views of assessment design, marking and feedback in higher education: a case for professionalisation? *Assessment & Evaluation in Higher Education*, 1-13. doi:10.1080/02602938.2019.1592110

Norton, L., Norton, B., & Shannon, L. (2013). Revitalising assessment design - what is holding new lecturers back? *Higher Education, 66*(2), 233-251. doi:10.1007/s

O'Brien, B., Harris, I., Beckman, T., Reed, D., & Cook, D. (2014). Standards for Reporting Qualitative Research: A Synthesis of Recommendations. *Academic Medicine, 89*(9), 1245-1251. doi:0.1097/ACM.0000000000000388

O'Conner, H., Madge, C., Shaw, R., & Wellens, J. (2008). Chapter 15: Internet-based interviewing *The SAGE handbook of research methods*.

O'Reilly, M., & Parker, N. (2012). 'Unsatisfactory Saturation': a critical exploration of the notion of saturated sample sizes in qualitative research. *Qualitative Research, 13*(2), 190-197. doi:10.1177/1468794112446106

Offerdahl, E., & Tomanek, D. (2011). Changes in instructors' assessment thinking related to experimentation with new strategies. *Assessment & Evaluation in Higher Education, 36*(7), 781-795. doi:10.1080/02602938.2010.488794

Ojo, E., & Booth, S. (2009). Internationalisation of higher education in a South African university: A phenomenographic study of students' conceptions. *Education as Change, 13*(2), 309-323. doi:10.1080/16823200903234851

Pang, M., & Ki, W. (2016). Revisiting the Idea of "Critical Aspects". *Scandinavian Journal of Educational Research, 60*(3), 323-336. doi:10.1080/00313831.2015.1119724

Pang, M., & Marton, F. (2013). Interaction between the learners' initial grasp of the object of learning and the learning resource afforded. *Instructional Science, 41*(6), 1065-1082.

Pedrosa-de-Jesus, M., & da Silva Lopes, B. (2011). The relationship between teaching and learning conceptions, preferred teaching approaches and questioning practices. *Research Papers in Education, 26*(2), 223-243. doi:10.1080/02671522.2011.561980

Pereira, D. (2016). Assessment in Higher Education and Quality of Learning: Perceptions, Practices and Implications. *Educational Studies, 43*(3).

Peshkin, A. (1993). The goodness of qualitative research. *Educational Researcher*. doi:https://doi.org/10.3102/0013189X022002023

Petty, N., Thomson, O., & Stew, G. (2012). Ready for a paradigm shift? Part 2: introducing qualitative research methodologies and methods. *Manuel Therapy, 17*(5), 378-384. doi:10.1016/j.math.2012.03.004

Pihl, E., Fridlund, B., & Martensson, J. (2011). Patients' experiences of physical limitations in daily life activities when suffering from chronic heart failure; a phenomenographic analysis. *Scandinavian Journal of Caring Science, 25*(1), 3-11. doi:10.1111/j.1471-6712.2010.00780.x

Pillow, W. (2003). Confession, catharsis, or cure? Rethinking the uses of reflexivity as methodological power in qualitative research. *International Journal of Qualitative Studies in Education, 16*(2), 175-196. doi:10.1080/0951839032000060635

Popham, W. (2009). Assessment Literacy for Teachers: Faddish or Fundamental? *Theory Into Practice, 48*(1), 4-11. doi:10.1080/00405840802577536

Postareff, L., Virtanen, V., Katajavuori, N., & Lindblom-Ylänne, S. (2012). Academics' conceptions of assessment and their assessment practices. *Studies in Educational Evaluation, 38*(3-4), 84-92. doi:10.1016/j.stueduc.2012.06.003

Prebble, T., Hargraves, H., Leach, L., Naidoo, K., Suddaby, G., & Zepke, N. (2004). Impact of Student Support Services and Academic Development Programmes on Student Outcomes in Undergraduate Tertiary Study - A Synthesis of the Research. *Massey University College of Education, Ministry of Education, New Zealand*, 1-141.

Price, M., Carroll, J., O'Donovan, B., & Rust, C. (2011). If I was going there I wouldn't start from here: a critical commentary on current assessment practice. *Assessment & Evaluation in Higher Education, 36*(4), 479-492. doi:10.1080/02602930903512883

Price, M., Handley, K., Millar, J., & O'Donovan, B. (2010). Feedback: all that effort, but what is the effect? *Assessment & Evaluation in Higher Education, 35*(3), 277-289. doi:10.1080/02602930903541007

Prideaux, D., & Bligh, J. (2002). Research in medical education - asking the right questions. *Medical Education, 36*(12), 1114-1115.

Probst, B. (2015). The Eye Regards Itself: Benefits and Challenges of Reflexivity in Qualitative Social Work Research. *Social Work Research, 39*(1), 37-48. doi:10.1093/swr/svu028

Pugh, D., Desjardins, I., & Eva, K. (2017). How do formative objective structured clinical examinations drive learning? Analysis of residents' perceptions. *Medical Teacher*, 1-8. doi:10.1080/0142159x.2017.1388502

Qu, S., & Dumay, J. (2011). The qualitative research interview. *Qualitative Research in Accounting & Management, 8*(3), 238-264. doi:10.1108/11766091111162070

Quesada-Serra, V., Rodríguez-Gómez, G., & Ibarra-Sáiz, M. S. (2016). What are we missing? Spanish lecturers' perceptions of their assessment practices. *Innovations in Education and Teaching International, 53*(1), 48-59. doi:10.1080/14703297.2014.930353

Ramani, S., & Mann, K. (2016). Introducing medical educators to qualitative study design: Twelve tips from inception to completion. *Medical Teacher, 38*(5), 456-463. doi:10.3109/0142159X.2015.1035244

Rear, D. (2018). One size fits all? The limitations of standardised assessment in critical thinking. *Assessment & Evaluation in Higher Education*, 1-12. doi:10.1080/02602938.2018.1526255

Reed, D., Cook, D., Beckman, T., Levine, R., Kern, D., & Wright, S. (2007). Association Between Funding and Quality of Published Medical Education Research. *Journal of the American Medical Association, 298*(9), 1002-1009.

Rees, C., & Knight, L. (2007). Viewpoint: The Trouble with Assessing Students' Professionalism: Theoretical Insights from Sociocognitive Psychology. *Academic Medicine, 82*(1), 46-50.

Rees, C., Knight, L., & Cleland, J. (2009). Medical educators' metaphoric talk about their assessment relationships with students: 'you don't want to sort of be the one who sticks

the knife in them'. *Assessment & Evaluation in Higher Education, 34*(4), 455-467. doi:10.1080/02602930802071098

Richardson, J. (1999). The Concepts and Methods of Phenomenographic Research. *Review of Educational Research, 69*(1), 53-82.

Ringsted, C., Hodges, B., & Scherpbier, A. (2011). 'The research compass': an introduction to research in medical education: AMEE Guide no. 56. *Medical Teacher, 33*(9), 695-709. doi:10.3109/0142159X.2011.595436

Roediger, H., & Karpicke, J. (2006a). The Power of Testing Memory: Basic Research and Implications for Educational Practice. *Perspectives on Pyschological Science, 1*(3), 181-210.

Roediger, H., & Karpicke, J. (2006b). Test-Enhanced Learning: Taking Memory Tests Improves Long-Term Retention. *Psychological Science, 17*(3), 249-255.

Rolin, K. (2009). Standpoint Theory as a Methodology for the Study of Power Relations. *Hypatia, 24*(4), 218-226.

Rosenblum, N., Kluijtmans, M., & ten Cate, O. (2016). Professional Identity Formation and the Clinician–Scientist. *Academic Medicine, 91*(12), 1612-1617. doi:10.1097/acm.0000000000001252

Ross, M. (2017). Conceptions of teaching: an illustrated review. *The Clinical Teacher, 14*, 8-14.

Ross, M., Green, S., Salisbury-Glennon, J., & Tollefson, N. (2006). College students' study strategies as a function of testing - an investigation into metacognitive self-regulation. *Innovative Higher Education, 30*(5), 361-375. doi:10.1007/s10755-005-9004-2

Ross, M., Salisbury-Glennon, J., Guarino, A., Reed, C., & Marshall, M. (2003). Situated Self-Regulation: Modeling the Interrelationships Among Instruction, Assessment, Learning Strategies and Academic Performance. *Educational Research and Evaluation, 9*(2), 189-209. doi:10.1076/edre.9.2.189.14211

Rotgans, J. (2011). The themes, institutions, and people of medical education research 1988-2010: content analysis of abstracts from six journals. *Advanced Health Science Education Theory and Practice, 17*(4), 515-527. doi:10.1007/s10459-011-9328-x

Rovio-Johansson, A., & Ingerman, A. (2016). Continuity and Development in the Phenomenography and Variation Theory Tradition. *Scandinavian Journal of Educational Research, 60*(3), 257-271. doi:10.1080/00313831.2016.1148074

Rushton, A. (2005). Formative assessment: a key to deep learning? *Medical Teacher, 27*(6), 509-513. doi:10.1080/01421590500129159

Samuelowicz, K., & Bain, J. (2002). Identifying academics' orientations to assessment practice. *Higher Education, 43*(2), 173-201.

Sandy, P. (2014). Factors Affecting Assessment of Student Nurses' Clinical Practice:A Phenomenographic Exploration of the Experiences and Understanding of Mentors of a Mental Health Service in England. *International Journal of Education Science, 7*(1), 57-66.

Santiago-Delefosse, M., Gavin, A., Bruchez, C., Roux, P., & Stephen, S. (2016). Quality of qualitative research in the health sciences: Analysis of the common criteria present in 58 assessment guidelines by expert users. *Social Science Medicine, 148*, 142-151. doi:10.1016/j.socscimed.2015.11.007

Schiekirka-Schwake, S., Anders, S., von Steinbüchel, N., Becker, J., & Raupach, T. (2017). Facilitators of high-quality teaching in medical school: findings from a nationwide survey among clinical teachers. *Medical Education, 17*, 178-186. doi:0.1186/s12909-017-1000-6

Schut, S., Driessen, E., van Tartwijk, J., van der Vleuten, C., & Heeneman, S. (2018). Stakes in the eye of the beholder: an international study of learners' perceptions within programmatic assessment. *Medical Education*. doi:10.1111/medu.13532

Schuwirth, L., & van der Vleuten, C. (2010). How to design a useful test: the principles and assessment. *Understanding Medical Education: Evidence, theory and practice. Association for the Study of Medical Education.*

Schuwirth, L., & van der Vleuten, C. (2011a). General overview of the theories used in assessment: AMEE Guide No. 57. *Medical Teacher, 33*(10), 783-797. doi:10.3109/0142159X.2011.611022

Schuwirth, L., & van der Vleuten, C. (2011b). Programmatic assessment: From assessment of learning to assessment for learning. *Medical Teacher, 33*(6), 478-485. doi:10.3109/0142159X.2011.565828

Segers, M., & Tillema, H. (2011). How do Dutch secondary teachers and students conceive the purpose of assessment? *Studies in Educational Evaluation, 37*(1), 49-54. doi:10.1016/j.stueduc.2011.03.008

Shay, S. (2004). The Assessment of Complex Performance: A Socially Situated Interpretive Act. *Harvard Educational Review, 74*(3), 307-329.

Shay, S., & Peseta, T. (2016). A socially just curriculum reform agenda. *Teaching in Higher Education, 21*(4), 361-366. doi:10.1080/13562517.2016.1159057

Sheeran, P. (2002). Intention—Behavior Relations: A Conceptual and Empirical Review. *European Review of Social Psychology, 12*(1), 1-36. doi:10.1080/14792772143000003

Shumway, J., & Harden, R. (2003). AMEE Guide No. 25: The assessment of learning outcomes for the competent and reflective physician. *Medical Teacher, 25*(6), 569-584. doi:10.1080/0142159032000151907

Signorini, P., Wiesemes, R., & Murphy, R. (2009). Developing alternative frameworks for exploring intercultural learning: a critique of Hofstede's cultural difference model. *Teaching in Higher Education, 14*(3), 253-264. doi:10.1080/13562510902898825

Sjostrom, B., & Dahlgren, L. (2002). Applying phenomenography in nursing research. *Journal of Advanced Nursing, 40*(3), 339-345.

Skidmore, S., Hsu, H., & Fuller, M. (2018). A person-centred approach to understanding cultures of assessment. *Assessment & Evaluation in Higher Education*, 1-17. doi:10.1080/02602938.2018.1447082

Smith, P. (2006). When elephants fight, the grass gets trampled: the GLOBE and Hofstede projects. *Journal of International Business Studies, 37*, 915-921.

Sonneman, U. (1954). Existence and therapy: an introduction to phenomenological psychology and existential analysis. *Grune & Stratton*, 372.

Sorinola, O., Thistlethwaite, J., Davies, D., & Peile, E. (2017). Realist evaluation of faculty development for medical educators What works for whom and why in the long term. *Medical Teacher, 39*(4), 422-429. doi:0.1080/0142159X.2017.1293238

Stalmeijer, R., McNaughton, N., & van Mook, W. (2014). Using focus groups in medical education research: AMEE Guide No. 91. *Medical Teacher, 36*(11), 923-939. doi:10.3109/0142159X.2014.917165

Steinert, Y., O'Sullivan, P., & Irby, D. (2019). Strengthening Teachers' Professional Identities Through Faculty Development. *Academic Medicine*. doi:10.1097/ACM.0000000000002695

Stenfors-Hayes, T., Hult, H., & Dahlgren, L. (2011). What does it mean to be a good teacher and clinical supervisor in medical education? *Advances in Health Science Education, Theory, and Practice, 16*(2), 197-210. doi:10.1007/s10459-010-9255-2

Stenfors-Hayes, T., Hult, H., & Dahlgren, L. (2012). Three ways of understanding development as a teacher. *European Journal of Dental Education, 16*(1), e151-157. doi:10.1111/j.1600-0579.2011.00690.x

Stenfors-Hayes, T., Hult, H., & Dahlgren, M. (2013). A phenomenographic approach to research in medical education. *Medical Education, 47*(3), 261-270. doi:10.1111/medu.12101

Stenfors-Hayes, T., Weurlander, M., Dahlgren, L., & Hult, H. (2010). Medical teachers' professional development – perceived barriers and opportunities. *Teaching in Higher Education, 15*(4), 399-408. doi:10.1080/13562517.2010.493352

Stenfors-Hayes, T., Hult, H., & Dahlgren, M. (2013). A phenomenographic approach to research in medical education. *Medical Education, 47*(3), 261-270. doi:0.1111/medu.12101

Stoetzler, M., & Yuval-Davis, N. (2002). Standpoint theory, situated knowledge and the situated imagination. *Feminist Theory, 3*(3), 315-333.

Stone, S., Ellers, B., Holmes, D., Orgren, R., Qualters, D., & Thompson, J. (2002). Identifying oneself as a teacher: the perceptions of preceptors. *Medical Education, 36*, 180-185.

Suhoyo, Y., Van Hell, E., Kerdijk, W., Emilia, O., Schonrock-Adema, J., Kuks, J., & Cohen-Schotanus, J. (2017). Influence of feedback characteristics on perceived learning value of feedback in clerkships: does culture matter? *BMC Medical Education, 17*(1), 69. doi:10.1186/s12909-017-0904-5

Suhoyo, Y., van Hell, E., Prihatiningsih, T., Kuks, J., & Cohen-Schotanus, J. (2014). Exploring cultural differences in feedback processes and perceived instructiveness during clerkships: replicating a Dutch study in Indonesia. *Medical Teacher, 36*(3), 223-229. doi:10.3109/0142159X.2013.853117

Swanwick, T. (2008). See one, do one, then what? Faculty development in postgraduate medical education. *Postgraduate Medical Journal, 84*(993), 339-343. doi:10.1136/pgmj.2008.068288

Tavakol, M., & Dennick, R. (2017). The foundations of measurement and assessment in medical education. *Medical Teacher*, 1-6. doi:10.1080/0142159X.2017.1359521

Tavakol, M., & Sandars, J. (2014a). Quantitative and qualitative methods in medical education research: AMEE Guide No 90: Part I. *Medical Teacher, 36*(9), 746-756. doi:10.3109/0142159X.2014.915298

Tavakol, M., & Sandars, J. (2014b). Quantitative and qualitative methods in medical education research: AMEE Guide No 90: Part II. *Medical Teacher, 36*(10), 838-848. doi:10.3109/0142159X.2014.915297

Taylor, D. (2011). 'They are using laptops, we are using boxes': township learners' conceptions of Expo. *African Journal of Research in Mathematics, Science and Technology Education, 15*(1), 67-79. doi:10.1080/10288457.2011.10740702

Telio, S., Ajjawi, R., & Regehr, G. (2015). The "educational alliance" as a framework for reconceptualizing feedback in medical education. *Academic Medicine, 90*(5), 609-614. doi:0.1097/ACM.0000000000000560

ten Cate, O., & Regehr, G. (2018). The Power of Subjectivity in the Assessment of Medical Trainees. *Academic Medicine*. doi:10.1097/acm.0000000000002495

Terry, D., Hogg, M., & White, K. (2010). The theory of planned behaviour: Self-identity, social identity and group norms. *British Journal of Social Psychology*.

Thanh, N., & Thanh, T. (2015). The Interconnection Between Interpretivist Paradigm and Qualitative Methods in Education. *American Journal of Educational Science, 1*(2), 24-27.

Tight, M. (2016). Phenomenography: the development and application of an innovative research design in higher education research. *International Journal of Social Research Methodology, 19*(3), 319-338. doi:10.1080/13645579.2015.1010284

Torre, D., Schuwirth, L., & van der Vleuten, C. (2019). Theoretical considerations on programmatic assessment. *Medical Teacher*, 1-8. doi:10.1080/0142159x.2019.1672863

Trigwell, K., Prosser, M., & Waterhouse, F. (1999). Relations between teachers' approaches to teaching and students' approaches to learning. *Higher Education, 37*(1), 57-70.

Tsai, S., Ho, M., Hirsh, D., & Kern, D. (2012). Defiance, compliance, or alliance? How we developed a medical professionalism curriculum that deliberately connects to cultural context. *Medical Teacher, 34*(8), 614-617. doi:10.3109/0142159X.2012.684913

Tung, R., & Verbeke, A. (2010). Beyond Hofstede and GLOBE- Improving the quality of cross-cultural research. *Journal of International Business Studies, 41*(8), 1259-1274.

Tutarel, O. (2002). Geographical distribution of publications in the field of medical education. *Medical Education, 2*(3).

van der Ridder, J., McGaghie, W., Stokking, K., & ten Cate, O. (2015). Variables that affect the process and outcome of feedback, relevant for medical training: a meta-review. *Medical Education, 49*(7), 658-673.

van der Vleuten, C., & Dannefer, E. (2012). Towards a systems approach to assessment. *Medical Teacher, 34*(3), 185-186. doi:10.3109/0142159X.2012.652240

van der Vleuten, C., & Driessen, E. (2014). What would happen to education if we take education evidence seriously? *Perspectives on Medical Education, 3*(3), 222-232. doi:10.1007/s40037-014-0129-9

van der Vleuten, C., Schuwirth, L., Driessen, E., Dijkstra, J., Tigelaar, D., Baartman, L., & van Tartwijk, J. (2012). A model for programmatic assessment fit for purpose. *Medical Teacher, 34*(3), 205-214. doi:10.3109/0142159X.2012.652239

van der Vleuten, C., Schuwirth, L., Driessen, E., Govaerts, M., & Heeneman, S. (2014). 12 Tips for programmatic assessment. *Medical Teacher*, 1-6. doi:10.3109/0142159X.2014.973388

van Etten, S., Freebern, G., & Pressley, M. (1997). College Students' Beliefs about Exam Preparation. *Contemporary Educational Psychology, 22*(2), 192-212.

van Lankveld, T., Schoonenboom, J., Volman, M., Croiset, G., & Beishuizen, J. (2016). Developing a teacher identity in the university context: a systematic review of the literature. *Higher*

Education Research & Development, 36(2), 325-342. doi:10.1080/07294360.2016.1208154

van Schalkwyk, S., Cilliers, F., Adendorff, H., Cattell, K., & Herman, N. (2013). Journeys of growth towards the professional learning of academics: understanding the role of educational development. *International Journal for Academic Development, 18*(2), 139-151. doi:10.1080/1360144X.2012.673490

van Schalkwyk, S., Leibowitz, B., Herman, N., & Farmer, J. (2015). Reflections on professional learning: Choices, context and culture. *Studies in Educational Evaluation, 46*, 4-10. doi:10.1016/j.stueduc.2015.03.002

Varela, O., & Premeaux, S. (2008). Do Cross-Cultural Values Affect Multisource Feedback Dynamics? The Case of High Power Distance and Collectivism in Two Latin American Countries. *International Journal of Selection and Assessment*

Varpio, L., Ajjawi, R., Monrouxe, L., O'Brien, B., & Rees, C. (2017). Shedding the cobra effect: problematising thematic emergence, triangulation, saturation and member checking. *Medical Education, 51*(1), 40-50. doi:10.1111/medu.13124

Verdonk, P., & Abma, T. (2013). Intersectionality and reflexivity in medical education research. *Medical Education, 47*(8), 754-756. doi:10.1111/medu.12258

Villarroel, V., Bloxham, S., Bruna, D., Bruna , C., & Herrera-Seda, C. (2017). Authentic assessment: creating a blueprint for course design. *Assessment & Evaluation in Higher Education*.

Visser-Wijnveen, G., van Driel, J., van der Rijst, R., Verloop, N., & Visser, A. (2009). The relationship between academics' conceptions of knowledge, research and teaching – a metaphor study. *Teaching in Higher Education, 14*(6), 673-686. doi:10.1080/13562510903315340

Wahlstrom, R., Lagerlov, P., Stalsby Lundborga, C., Veningac, C., Hummers-Pradierd, E., Dahlgren, L., & Denig, P. (2001). Variations in general practitioners' views of asthma management in four European countries. *Social Science and Medicine, 53*, 507-518.

Wald, H. (2015). Professional Identity (Trans)Formation in Medical Education. *Academic Medicine, 90*(6), 701-706. doi:10.1097/acm.0000000000000731

Walubo, A., Burch, V., Parmar, P., Raidoo, D., Cassimjee, M., Onia, R., & Ofei, F. (2003). A model for selecting assessment methods for evaluating medical students in African medical schools. *Academic Medicine, 78*(9), 899-906.

Wass, V., van der Vleuten, C., Shatzer, J., & Jones, R. (2001). Assessment of clinical competence. *The Lancet, 357*(9260), 945-949. doi:10.1016/s0140-6736(00)04221-5

Watkins, D., Dahlin, B., & Ekholm, M. (2005). Awareness of the backwash effect of assessment-A phenomenographic study of the views of Hong Kong and Swedish lecturers. *Instructional Science, 33*(4), 283-309.

Watling, C., & Ginsburg, S. (2018). Assessment, feedback and the alchemy of learning. *Medical Education*. doi:10.1111/medu.13645

Webb, G. (1997a). Contesting learning theory: a response to Entwistle and Ekeblad. *Higher Education, 33*, 225-230.

Webb, G. (1997b). Deconstructing deep and surface: Towards a critique of phenomenography. *Higher Education, 33*(195-212).

Weller, S. (2017). Using internet video calls in qualitative (longitudinal) interviews: some implications for rapport. *International Journal of Social Research Methodology, 20*(6), 613-625. doi:10.1080/13645579.2016.1269505

Whiting, L. (2008). Semi-structured interviews: guidance for novice researchers. *Nursing Standard, 22*(23).

WHO. (2018). World health statistics 2018: Monitoring health for the sustainable development goals.

Wilbur, K., Bacha, R., & Abdelaziz, S. (2017). How does culture affect experiential training feedback in exported Canadian health professional curricula? *International Journal of Medical Education., 8*, 91-98.

Wilhelmsson, N., Dahlgren, L., Hult, H., & Josephson, A. (2011). On the anatomy of understanding. *Studies in Higher Education, 36*(2), 153-165. doi:10.1080/03075070903514054

Witman, P., & Richlin, L. (2007). The Status of the Scholarship of Teaching and Learning in the Disciplines. *International Journal for the Scholarship of Teaching and Learning, 1*(1), 1-17. doi:10.20429/ijsotl.2007.010114

Woollacott, L., Booth, S., & Cameron, A. (2013). Knowing your students in large diverse classes: a phenomenographic case study. *Higher Education, 67*(6), 747-760. doi:10.1007/s10734-013-9664-2

Wu, Q., & Jessop, T. (2018). Formative assessment: missing in action in both research-intensive and teaching focused universities? *Assessment & Evaluation in Higher Education, 43*(7), 1019-1031. doi:10.1080/02602938.2018.1426097

Yan Fung, M. (2005). A Philosophy of Teaching Practicum: construction of a personal theory of teaching and learning. *Teacher Development, 9*(1).